RF and Microwave Measurement Fundamentals

H7215A#101 v2.0 — January 2001

Global Education R&D

Knowledge Services Organization

Agilent Technologies

RF and Microwave Measurement Fundamentals

H7215A#101 v2.0 — January 2001

Global Education R&D

Knowledge Services Organization

Spark your educational insight @

www.agilent.com/find/tmeducation

800 593 6632

Agilent Technologies

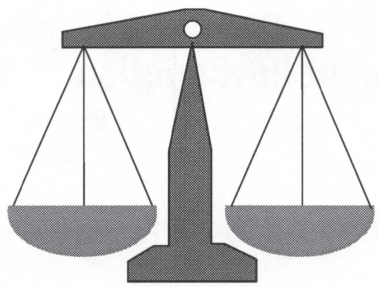
RF & Microwave Measurement Fundamentals

1 — Measurement Introduction

Agilent Technologies

A Definition

- **Measurement**
  - **The process of converting our physical world into meaningful pieces of information.**



An Early Instrument

01 – Measurement Introduction
H7215A#101 v2.0

Agilent Technologies

Page 2

Measurements and the resulting data allow us to improve our daily processes.

Measurement Terms

- Physical property or condition to be measured
  - measurand
- Converts energy of one form to another, typically to an electrical signal
  - transducer
- Tool to measure with
  - instrument

MEASURANDS

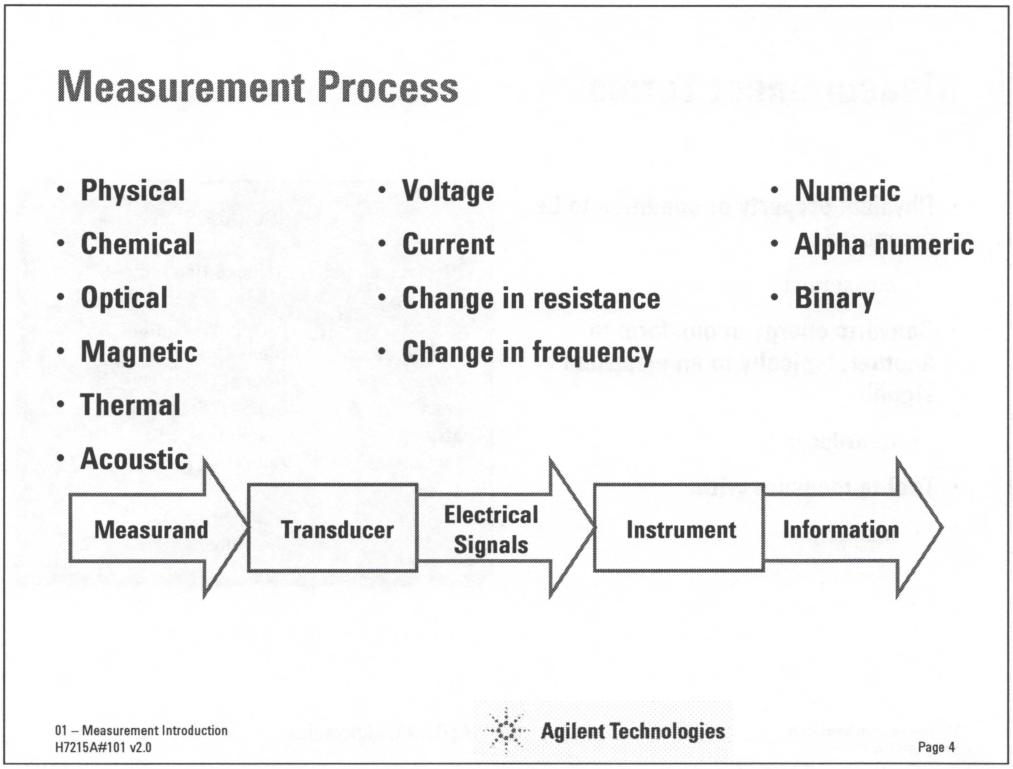
Displacements	Relative Humidity
Position	Atomic & Surface Profiles
Velocity	Gas Concentration
Acceleration	Blood Gases, pH
Force	Optical Fields
Load	Infrared Radiation
Strain	Magnetic Fields
Torque	Acoustic Fields
Vibrations	Flow
Temperature	Rotation
Pressure	Vacuum

01 – Measurement Introduction
H7215A#101 v2.0

Agilent Technologies

Page 3

We have to have something to measure, and a way to measure it.



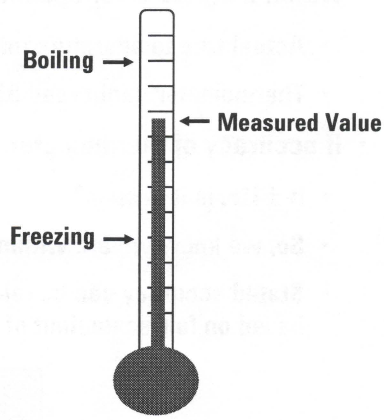
The transducer does the conversion to an electrical signal, and the instrument does the actual measuring.

Examples of instruments that measure an electrical signal are a voltmeter, ammeter, ohmmeter, oscilloscope for the time domain, and a spectrum analyzer for the frequency domain.

A Basic Measurement

- **Instrument is calibrated to a standard.**
- **A measured value is compared against a standard value.**

Analog Thermometer



Typically, the manufacturer does the calibration. Some instruments require recalibration over time and the user may or may not perform it.

Instrument Performance - Accuracy

- **Accuracy is the ability of an instrument to measure the actual value within a stated error specification.**
  - **Actual true temperature might be 50°F**
  - **Thermometer might read 51°F**
- **If accuracy of thermometer is stated to ±1%, is thermometer in spec?**
  - **If ±1°F, is it in spec?**
  - **So, we know we are within ±1% or ±1°F of the actual true temperature.**
  - **Stated accuracy can be relative to the value measured (±1%) or absolute based on full-scale limit of instrument (±1°F).**

ACCURACY

01 – Measurement Introduction
H7215A#101 v2.0

Agilent Technologies

Page 6

Instrument Performance - Measurement Uncertainty

- **The uncertainty of the result of a measurement reflects the lack of exact knowledge of the value of the measurand (true value).**
  - **Remember, the result of measurement even after correction is an estimate of the value of the measurand.**
- **The result of a measurement is complete only when accompanied by a statement of uncertainty.**
  - **These statements were paraphrased from "Guide to the Expression of Uncertainty in Measurement" ISO 1995**

UNCERTAINTY

01 – Measurement Introduction
H7215A#101 v2.0

Agilent Technologies

Page 7

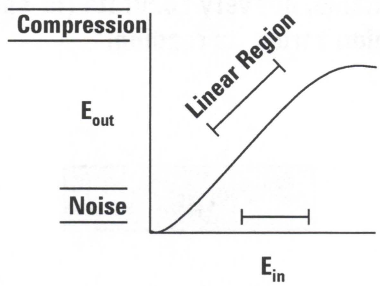
Instrument Performance - Resolution

- Resolution is the smallest change in the property being measured that an instrument can detect.
- If the actual temperature rose 0.1°F, but the thermometer still read 51°F, then that thermometer could not resolve a 0.1° change.
- The user might be able to resolve 0.5°F, depending on how long the thermometer was and what range of temperature it could measure.

RESOLUTION

Instrument Performance - Linearity

- **An instrument is linear if it responds uniformly to equal changes in the measured property.**



01 – Measurement Introduction
H7215A#101 v2.0
Agilent Technologies
Page 9

The noise region of the curve is where the input signal level is affected by spurious signals that are about the same level.

The compression region of the curve is where any further increase in the input signal does not cause any change in the output signal.

Instrument Performance - Stability

- An instrument's ability to make repeatable measurements of an exact value of the measured property over time.
- What if the thermometer read 51°F one time, and 52°F the next time. It would not be very stable, not very accurate (in-spec, then out-of-spec), and you wouldn't trust its reading.

STABILITY

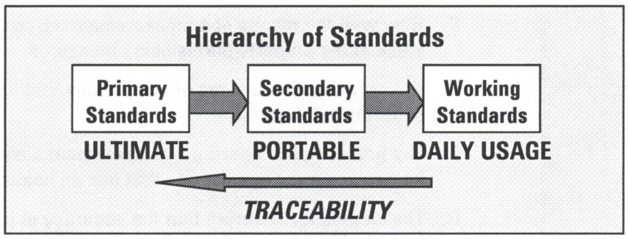
01 – Measurement Introduction
H7215A#101 v2.0

Agilent Technologies

Page 10

Instrument Performance - Standards

- **An acknowledged measure of comparison.**
- **An instrument's variation from a standard during calibration defines its accuracy.**



Standards must have traceability

STANDARDS

01 – Measurement Introduction
H7215A#101 v2.0

Agilent Technologies

Page 11

Governments typically define standards. NIST (National Institute of Standards and Technology) was formerly the NBS (National Bureau of Standards).

Exercise

1. Accuracy
2. Resolution
3. Linearity
4. Stability
5. Standard
6. Calibration
7. Traceability

A. The smallest change in a physical property that an instrument can sense.

B. The ability of an instrument to make repeatable measurements under exactly the same physical stimuli over time.

C. How well the results of a measurement agree with the actual value of the property that is being measured.

D. Acknowledged measures of comparison that are used to calibrate instruments.

E. The process of comparing a measurement device which has a known accuracy against one that has an unknown accuracy.

F. The method for ensuring that the accuracy of instruments can be linked to an accepted measurement reference source, or standard.

G. How constant the changes in response are, as a function of equal changes in the measured property.

01 – Measurement Introduction
H7215A#101 v2.0

Agilent Technologies

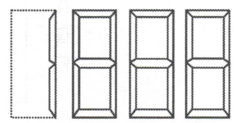
Page 12

Matching quiz.

Measurement Results - Types

- **Alphanumeric (ASCII data)**
- **Binary (computer data)**
- **Graphical (formatted to user)**
- **Numeric (meter reading)**

The 1/2 Digit



999 3-digit
1999 3-1/2 digit

01 – Measurement Introduction
H7215A#101 v2.0

Agilent Technologies

Page 13

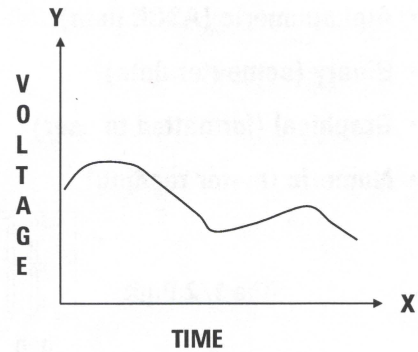
The extra 1/2 digit gives an instrument the ability to measure 100% above the normal full-scale range.

Binary data is usually provided by an instrument in the form of an interface circuit that can communicate to other electronic devices.

Graphical data is normally formatted by the instrument for displaying to the user.

Measurement Results - Graphical

- One piece of information per plot.
- Independent variable, usually time or frequency.



GRAPHICAL
(Cartesian coordinates)

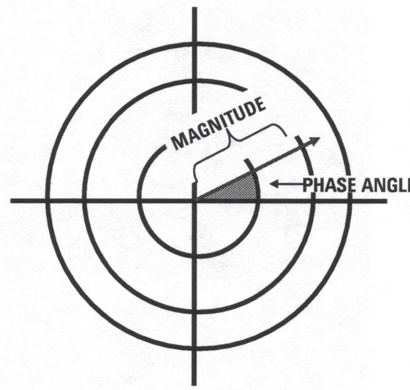
01 – Measurement Introduction
H7215A#101 v2.0

Agilent Technologies

Page 14

Measurement Results - Graphical Polar

- **Two pieces of information per plot.**
- **Independent variable, usually time or frequency.**



GRAPHICAL
(Polar coordinates)

01 – Measurement Introduction
H7215A#101 v2.0

Agilent Technologies

Page 15

You can easily convert between polar and Cartesian coordinates.

$$X = \text{magnitude} \times \cos(\theta)$$
$$Y = \text{magnitude} \times \sin(\theta)$$

$$\text{magnitude} = \text{sqrt}\,(X^2 + Y^2)$$
$$\text{angle} = \arctan\,(Y/X)$$

Measurement Units

Degrees (temperature)	Volts
Newtons / Meter^2 (pressure)	Amperes
Meters (distance)	Ohms
Meters / Second (velocity)	Farads
Meters / Second / Second (acceleration)	Henrys
Newton - meters (force)	Watts
Grams (mass)	Seconds
Joules (energy)	Hertz
Hogsheads (volume)	Degrees (phase)
Furlongs (distance)	Decibels (logarithmic)

01 – Measurement Introduction
H7215A#101 v2.0

Agilent Technologies

Page 16

Scaling Factors

Multiple	Prefix	Symbol
10^{18}	exa	E
10^{15}	peta	P
10^{12}	tera	T
10^{9}	giga	G
10^{6}	mega	M
10^{3}	kilo	k
10^{2}	hecto	h
10^{1}	deka	da
10^{-1}	deci	d
10^{-2}	centi	c
10^{-3}	milli	m
10^{-6}	micro	µ
10^{-9}	nano	n
10^{-12}	pico	p
10^{-15}	femto	f
10^{-18}	atto	a

01 – Measurement Introduction
H7215A#101 v2.0
Agilent Technologies
Page 17

Measurement Errors

- Actual values* (the value of the measurand)
- Measured values
- Errors terms (difference between measured and actual values)
  - Random (electrical noise, temperature changes)
  - Systematic (bad or no calibration, poor measurement methodology)
    - * The actual value is not knowable, but we will use the term anyway.

01 – Measurement Introduction
H7215A#101 v2.0

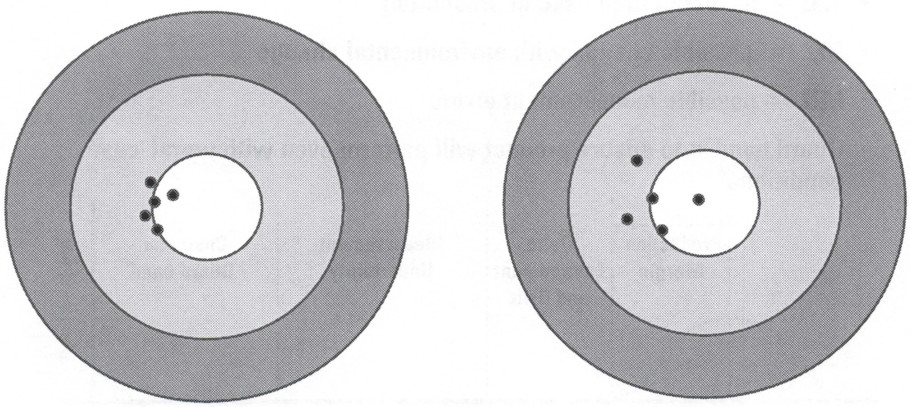
Agilent Technologies

Page 18

We usually focus on removing systematic errors from our measurements because we can do something about them. They are predictable and repeatable.

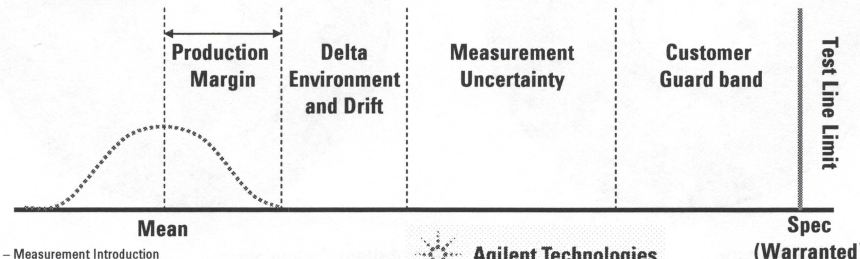
Random and Systematic Errors - Target Analogy

- **Both these marksmen have about the same systematic error but different random errors**



Measurements and Specifications

- For Agilent Technologies Instruments
  - Mean = average product performance
  - TLL = pass/fail limit used in production
  - DE = possible change with environmental change
  - MU = possible measurement errors
  - Guard band = to ensure product will perform even with worst-case conditions



RF & Microwave Measurement Fundamentals

2 — Transmission Lines

Agilent Technologies

Why A Special Class for Microwaves?

- The physics does not change but techniques change because:
  - The relationship to device size to wavelength
  - The relationship of path length to risetime
- And measurement uncertainties can be large

02 – Transmission Lines
H7215A#101 v2.0

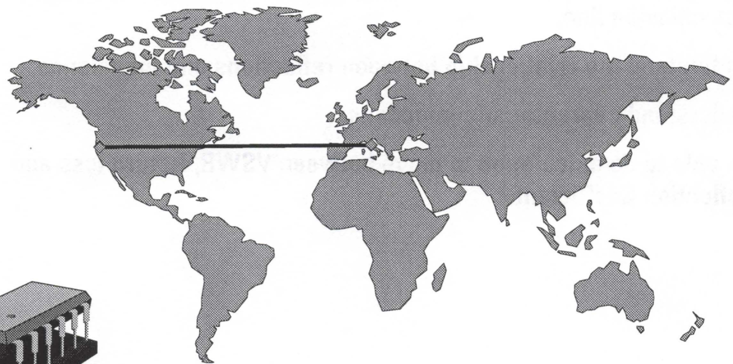
Agilent Technologies

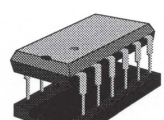
Page 2

Transmission lines are electrical connections which are used to efficiently transmit energy or information between a source and receiver. The distance between source and receiver may only be a few millimeters as across a chip in some microcircuit, or may be thousands of miles as in a trans oceanic submarine cable. In fact the concept of the transmission line was developed in the mid nineteenth century when developing long distance telegraphs.

Transmission Lines

- **A transmission line is a channel for the transmission of energy with minimal loss and distortion.**





02 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 3

Transmission Lines

In this lesson you will be introduced (or re-introduced) to the concept and some basic theory of transmission lines. The mathematics will be kept to a minimum although we make no apology for the occasional formula. Your instructor will describe the terms as they arise.

The concept of transmission lines is basic to the understanding of RF and microwave technology, although the concept is by no means limited to those high frequencies but is also applied at power line frequencies (50 or 60Hz.) for understanding the transmission of electrical power over long distances.

Objectives

- **At the completion of this class, you will:**
  - **Understand the need to move from simple "hookup" to the concept of the transmission line.**
  - **Understand the relationship between reflections and impedance.**
  - **Understand characteristic impedance.**
  - **Be able to do calculation to move between VSWR, Return loss and Reflection Coefficient.**

Wavelength and Frequency

• Velocity(3E8m/s) = frequency(Hz) × wavelength(m)

Kind of Wave	Frequency Hz	Wavelength (λ) m
Household AC	60/50	5000k/6000k
FM Radio	100M	3
Microwaves	1G/300G	0.3/0.001
Infrared	1E13	30000n
Visible Light	5E14	600n
X rays	3E18	0.1n

02 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 5

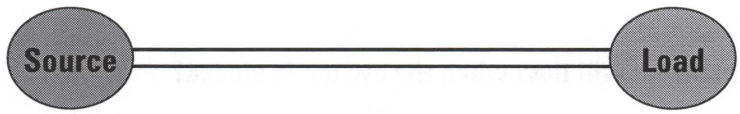
Wavelength and Frequency

Although the title of this course has RF and microwave as a description the concept of transmission lines can be applied at any frequency . For example the basis of this part of the class could apply to power engineering because although the frequency is certainly not microwave of RF the size of the transmission circuits are very long; a 300 mile transmission would represent about 10% of a wavelength at 60Hz.

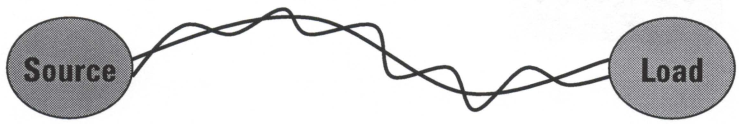
Transmission Line Agenda

- Characteristic Impedance
- Terminating a Transmission Line
- Reflections
- Impedance
  - What is the Smith Chart ?
- Standing Waves
- Some practical Transmission Lines

Transmission Lines versus Hookups



- Transmission Line
  - Uniform
  - Wavelength (delay) comparable or < with Size (risetime)
  - Voltages and current are Local (components are distributed)



- Non TL
  - Probably not uniform
  - Wavelength >> Size
  - Voltage and current Global (components are lumped)

02 – Transmission Lines
H7215A#101 v2.0

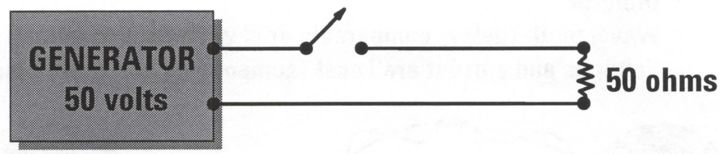
Agilent Technologies

Page 7

When we first learn about circuits, voltages and currents, the loops and branches contain lumped components, and the loops and branches may be treated discretely. For example a current path would exist from the source to the load and back again. With transmission lines, Ohm's law and Kirchoff's laws are still valid from the local point of view only.

Transmission Line – A Simple Electrical Circuit

- **What current will flow when the switch is closed?**



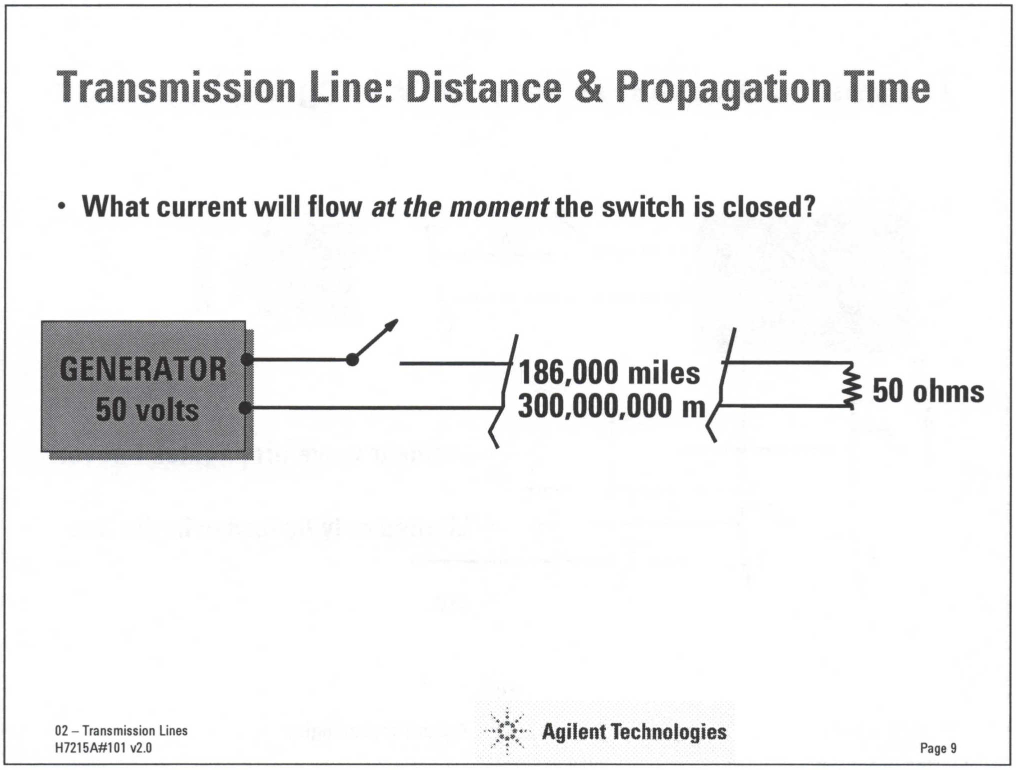
02 – Transmission Lines
H7215A#101 v2.0

$V=IR$

Agilent Technologies

Page 8

When the switch is closed a current of 1 Amp will flow. This result is based on Ohm's law and will be the basis for our understanding of transmission lines.



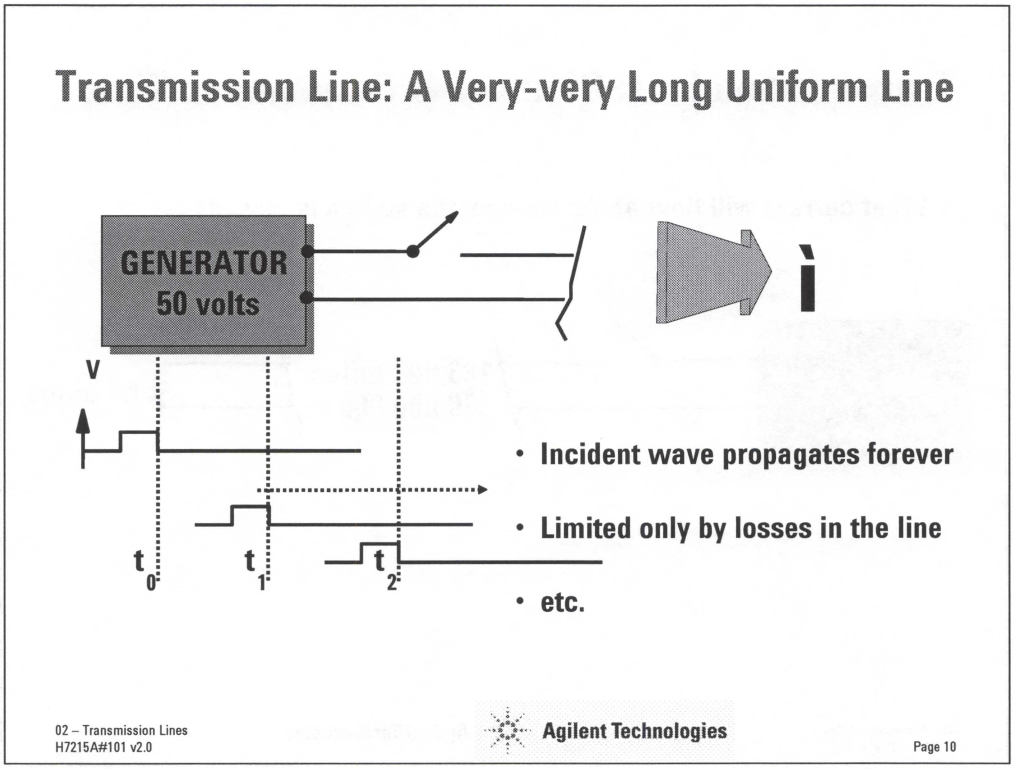
The difference between this and the preceding picture is only one of perception, to force us to think about what happens at the moment and a few moments after the switch is closed.

No information can travel faster than the speed of light, this speed is very close to 186,000 miles/second, or 300,000,000 m/second.

Even if the information about the switch closure moved with the speed of light, the generator would still supply the one amp due to the load, but this condition could only be true at a minimum of two seconds after switch closure. What happens up to this time will be the subject of the following presentation.

The following assumptions are made:

- The line is lossless
- The line is uniform

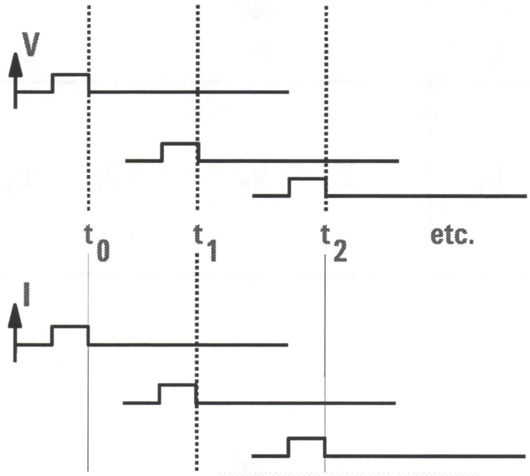


Imagine that this uniform lossless line is infinitely long. By uniform we mean it looks the same wherever it is cut, the same conductor size, the same material etc. If the switch closed then opened, causing a pulse to be created. We can think of this pulse travelling down the line forever.

This is the *incident* signal or *incident wave.*

Propagation in an Infinite Uniform Transmission Line

- **What current will be associated with the propagating voltage wave?**

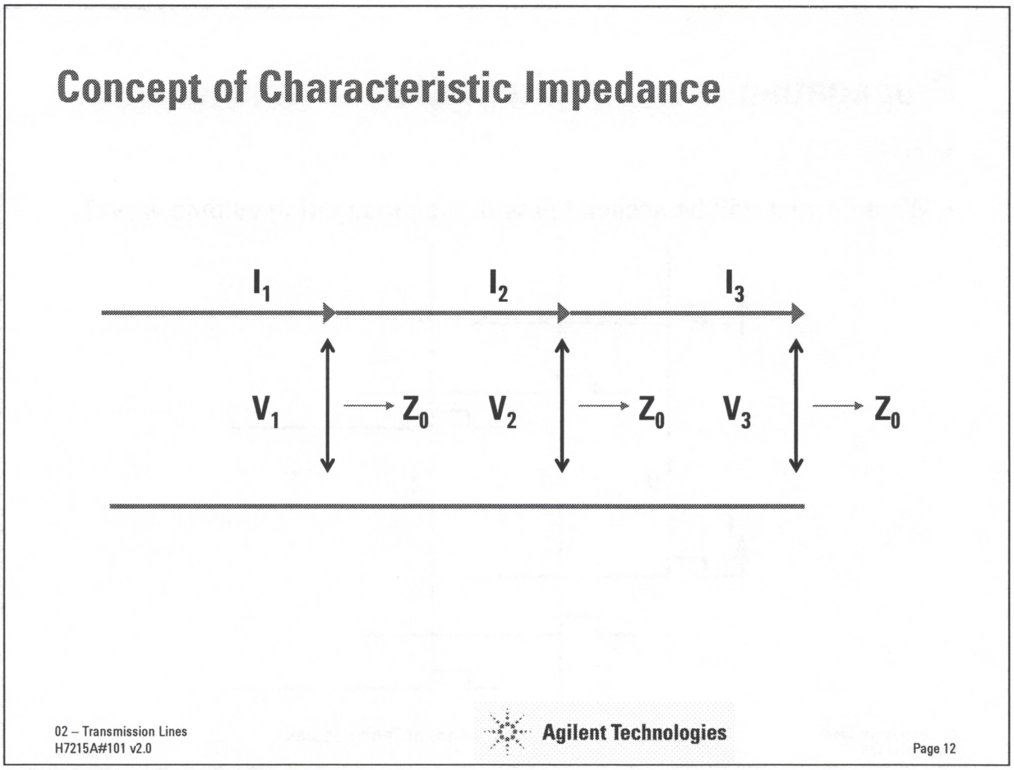


02 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 11

When the voltage pulse is formed an associated pulse of current is also formed. The usual explanation of current is in terms of electron flow, electrons are involved with the current wave but the drift speed of electrons is quite small (meters/sec) the key is the word *wave.* Waves propagate energy, without any physical displacement of, in this case, the electrons themselves over the length of the line.



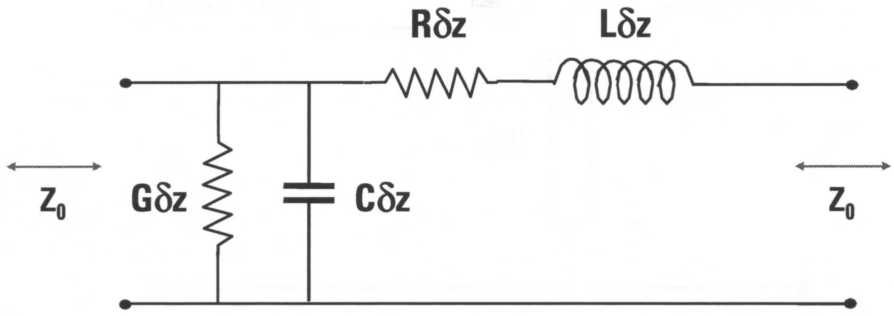
Out assumption is still that the line is infinitely long and uniform. Instead of a pulse, imagine that the generator is a sinewave source of a given frequency; this wave will continue its propagation for ever. If it were possible to measure the voltage and current at different points along the line, the ratio of the voltage to the current at these points would be constant. This ratio is an impedance, and this impedance is called the *characteristic impedance.*

This impedance exists without any knowledge of a load on the line. The line being infinitely long, this impedance must be related only to the physical properties of the line, the size of the wires, their distance apart, the nature of the substance between the wires.

These properties affect the capacitance, inductance , resistance, and leakage per unit length of the line. Of course real lines are never lossless but it's a useful approximation to make for many high frequency transmission lines.

Note: On looking at the diagram above, it appears that Kirchoff's law is being violated. I_1 is not I_2 so how is the equation balanced? The answer is to be found in Maxwell's equations which show the existence of the *displacement current* which will be normal to the conducting surfaces.

Equivalent Network to Determine the Characteristic Impedance

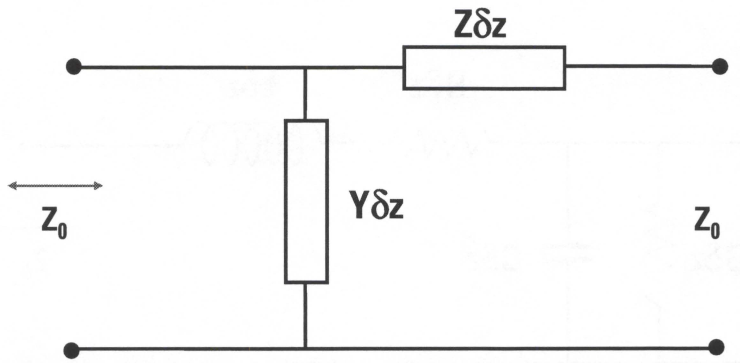


When analyzing a uniform transmission line's properties a model is used which is based on measurable constants:

- R - Resistance per meter - ohms/m
- C - Capacitance per meter - farads/m
- G - Conductance per meter - Seimens/m or mhos/m
- L - Inductance per meter - henrys/m

This model only assumes that the transmission line is uniform but not necessarily lossless as the presence of R and G testify.

Characteristic Impedance



02 – Transmission Lines
H7215A#101 v2.0
Agilent Technologies
Page 14

A non rigorous proof of the characteristic impedance equation:

Writing the input impedance in terms of the series and shunt components with a Z_0 termination:

$$Z_0 = \frac{\{Z_0 + (Z\delta z)\}\{1/(Y\delta z)\}}{Z_0 + (Z\delta z) + 1/(Y\delta z)}$$

Multiply numerator and denominator by $Y\delta z$ and treat the product $Z\delta z$. $Y\delta z$ as negligible.

$$= \frac{Z_0 + (Z\delta z)}{1 + Z_0(Y\delta z)}$$

$$\text{Solving for } Z_0$$

$$= \sqrt{\frac{Z}{Y}} = \sqrt{\frac{R + j\omega L}{G + j\omega C}}$$

Characteristic Impedance Z_0

$$Z_0 = \sqrt{\frac{R + j\omega L}{G + j\omega C}}$$

02 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 15

This fundamental formula, expresses an impedance for the line in terms of

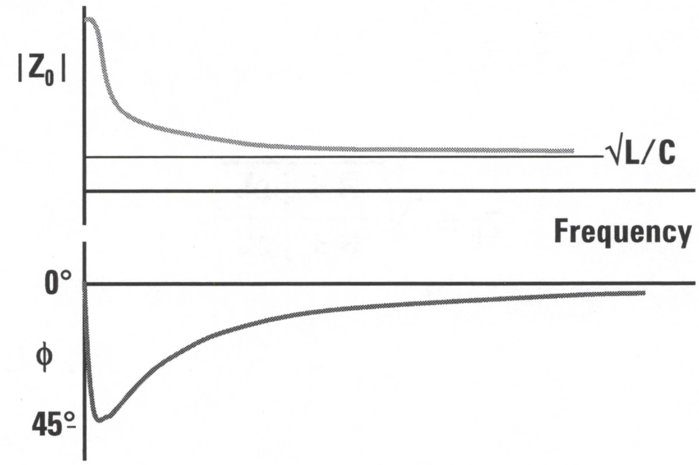
- R - A resistive loss along the length of the line.
- G - The conductance of the insulation between the conductors.
- L - The inductance along the conductors.
- C - The capacitance between the conductors.

All these parameters would be expressed as a per-unit-length basis.

The parameters G and R are properties of the materials and would be verified by material measurement.

The parameters L and C however can be derived from physical measurements, distances and dimensions of the conductors.

Variation of Characteristic Impedance with Frequency



02 – Transmission Lines
H7215A#101 v2.0

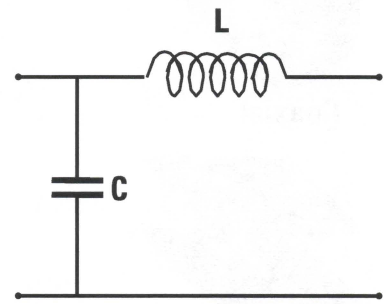
Agilent Technologies

Page 16

The Lossless Line

- **At high frequencies $\omega L >> R$ and $\omega C >> G$**

$$Z_0 = \sqrt{\frac{L}{C}}$$



02 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 17

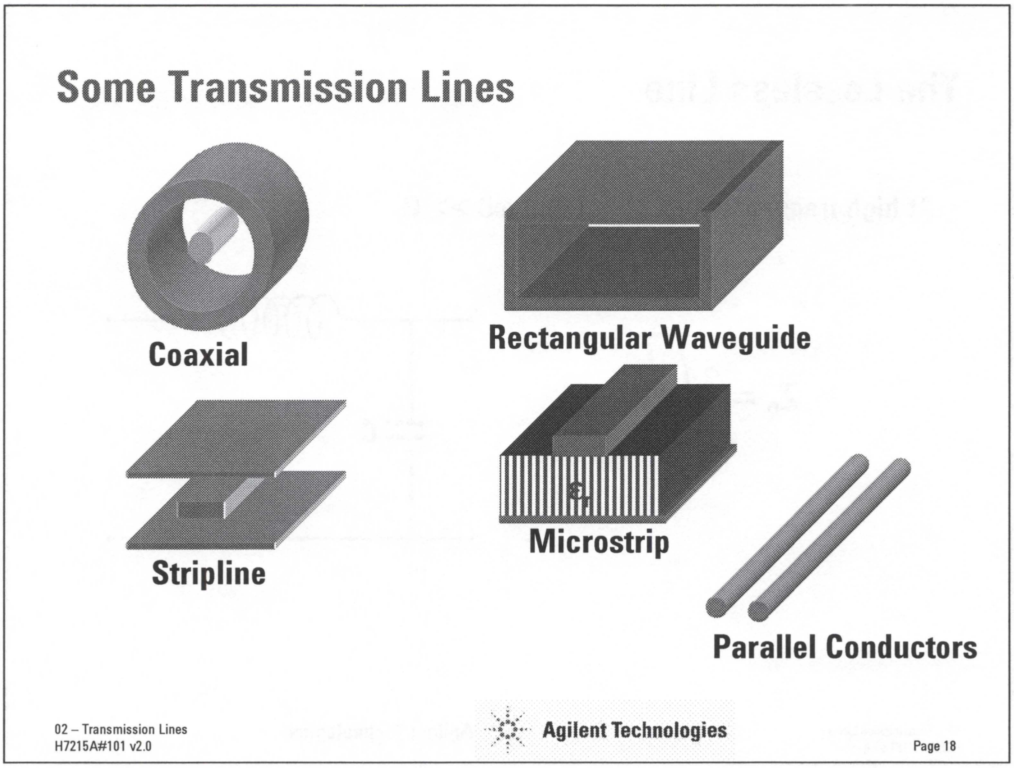
If G and R are negligible then the formula becomes very simple, this is the result for a lossless transmission line.

This result for a lossless line is remarkable for what is *not* in the formula.

- There is no "j". This operator shows that voltage and current are "out of phase"
- There is no "ω". This is the radian frequency of the wave = $2\pi f$

The lossless transmission line behaves from the point of view of the generator as a resistor. In the terms of mathematics the result is called *real.* In electrical engineering this means that current and voltage are in phase, and the impedance of this infinite lossless line looks like a resistor. And, also like an ideal resistor, is insensitive to frequency.

In dealing with high RF frequencies (Remember $\omega L >> R$ and $\omega C >> G$ because ω is large.) we often assume that the characteristic impedance is real. Typical examples are the 50 ohm and 75 ohm values found in RADAR and communications. Or 300 ohms found in the parallel wire antenna feeds for VHF TV.



Transmission Lines come in many forms, here are some well known types.

Velocity of Propagation

- **A single frequency sinewave will travel at a constant velocity, known as the phase velocity**
  - **A sinewave has no bandwidth**
- **Information signals. have bandwidth, a narrow part of this bandwidth is called a group.**
- **Information, the collection of groups, propagates at group velocities.**
- **Phase and group velocities assumed the same only for a lossless line.**
- **Phase and group velocities are different in a waveguide passband.**
- **The velocity in a medium compared to velocity in vacuum is the velocity factor.**
- **The velocity factor in a teflon insulated coax is about 0.6**

02 – Transmission Lines
H7215A#101 v2.0
Agilent Technologies
Page 19

The propagation of signals in a transmission line has two aspects:

- The phase shift per unit length. (which is related to velocity)
- The loss per unit length

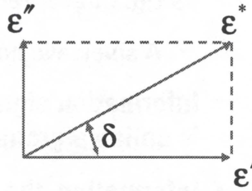
These are related to the primary constants of the line. For coaxial RF lines, since $R \ll \omega L$ and $G \ll \omega C$, the phase and group velocities can be considered the same and usually are some fraction of the velocity of light: 0.65 for Teflon (PTFE) filled coax. This factor is called the *velocity factor*.

For waveguide, phase velocity is greater than light velocity; group velocity is less than light velocity. $c = \sqrt{(v_p v_g)}$

Notes: Materials 1

- **Electric fields: Permittivity ε (dielectric constant κ)**

$$\varepsilon^* = \varepsilon' + j\,\varepsilon''$$



ε' is the storage part
ε'' is the dissipative part

$$\kappa = \varepsilon_r = \frac{\varepsilon'}{\varepsilon_0} \quad (\varepsilon'' \textit{ is assumed negligible})$$

$\tan\delta$ = loss tangent

ε_0 = Permitivity of free space = 8.854×10^{-12} Farad/m

Teflon $\kappa = 2.2$: Velocity factor $\frac{1}{\sqrt{\kappa}} = 0.674$

02 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 20

Notes: Materials 2

- **Magnetic fields: Permeability μ**

$$\mu^* = \mu' + j\,\mu''$$

μ' is the storage part
μ'' is the dissipative part

02 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 21

Notes: Skin Effect

- For AC the inductance increases toward the center of the conductor
- At high frequencies this means that current is confined to the surface (skin) of a conductor.
- Skin depth δ is the thickness of a layer where the current density drops to 1/e the value at the surface.
- $\delta = 5033 \sqrt{(\rho/\mu f)}$: ρ = resistivity ohms/cm3: μ = permeability
- Example: For copper $\delta(cm) = 6.62/\sqrt{f}$: f in Hz

02 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 22

Transmission Line Agenda

- Characteristic Impedance
- Terminating a Transmission Line
- Reflections
- Impedance
  - What is the Smith Chart ?
- Standing Waves
- Some practical Transmission Lines

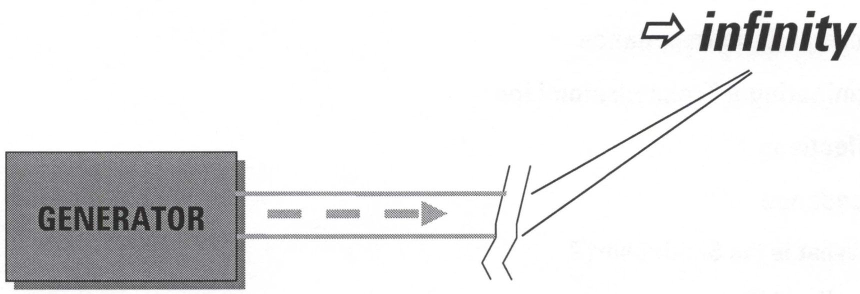
02 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 23

This is a short lesson about the loads that may be connected to a transmission line.

Infinite Transmission Line



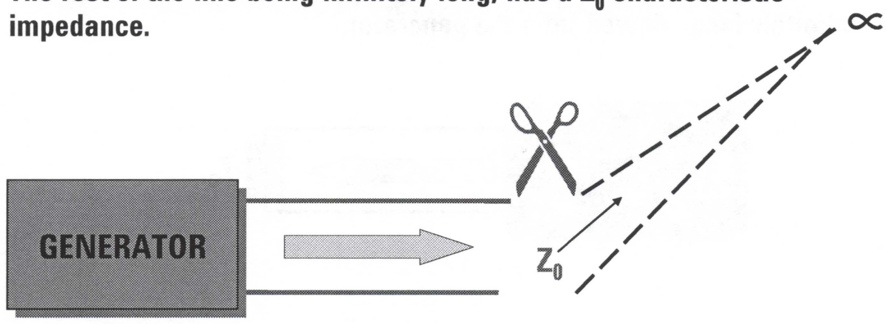
- Incident wave propagates forever
- Limited only by losses

02 – Transmission Lines
H7215A#101 v2.0
Agilent Technologies
Page 24

Imagine that the incident wave propagates forever, the relation between the voltage and current described by Z_0.

Terminating a Transmission Line

- **The rest of the line being infinitely long, has a Z_0 characteristic impedance.**

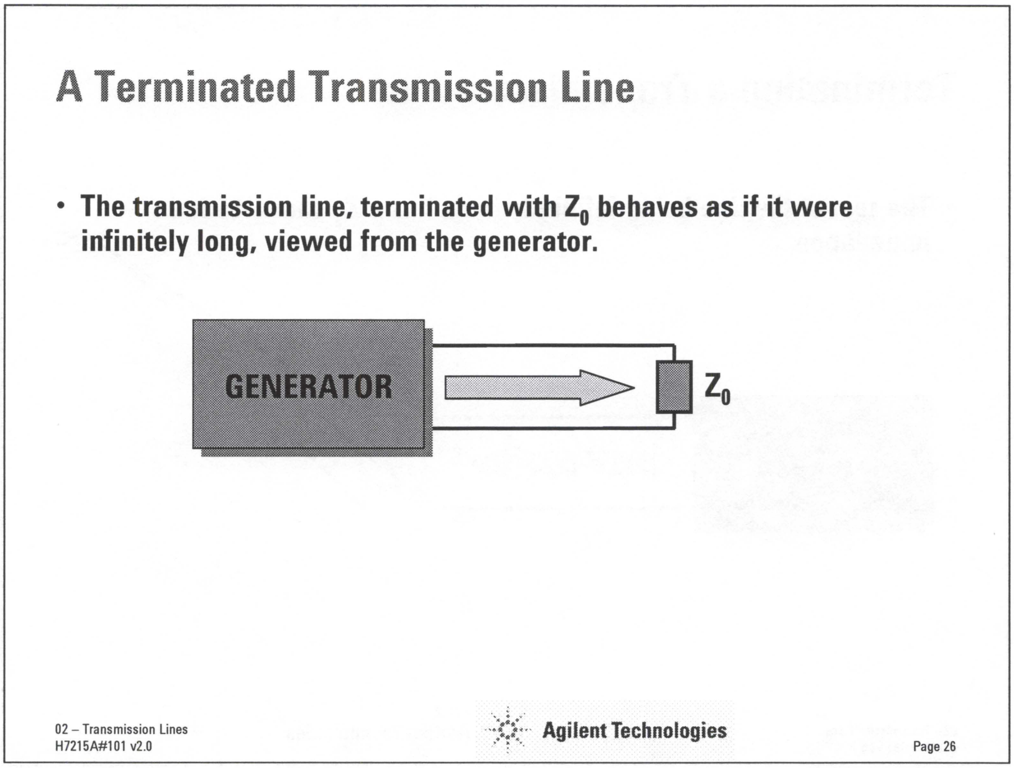


02 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 25

This is a simple idea, but it is an important one for the understanding of transmission lines. The imaginary infinitely long line has been cut off leaving some arbitrary length, the part we have imagined to be cut off would still be infinite in length, having a characteristic impedance of Z_0.



The length of transmission line if terminated with a lumped impedance with the same value as the characteristic impedance will appear to the generator as an infinite line. In other words the incident wave is propagating and is the only wave existing.

A Terminated Transmission Line has ...

- An incident wave
  - An impedance related to its primary constants; R, G, L & C
  - This impedance is called Characteristic Impedance

02 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

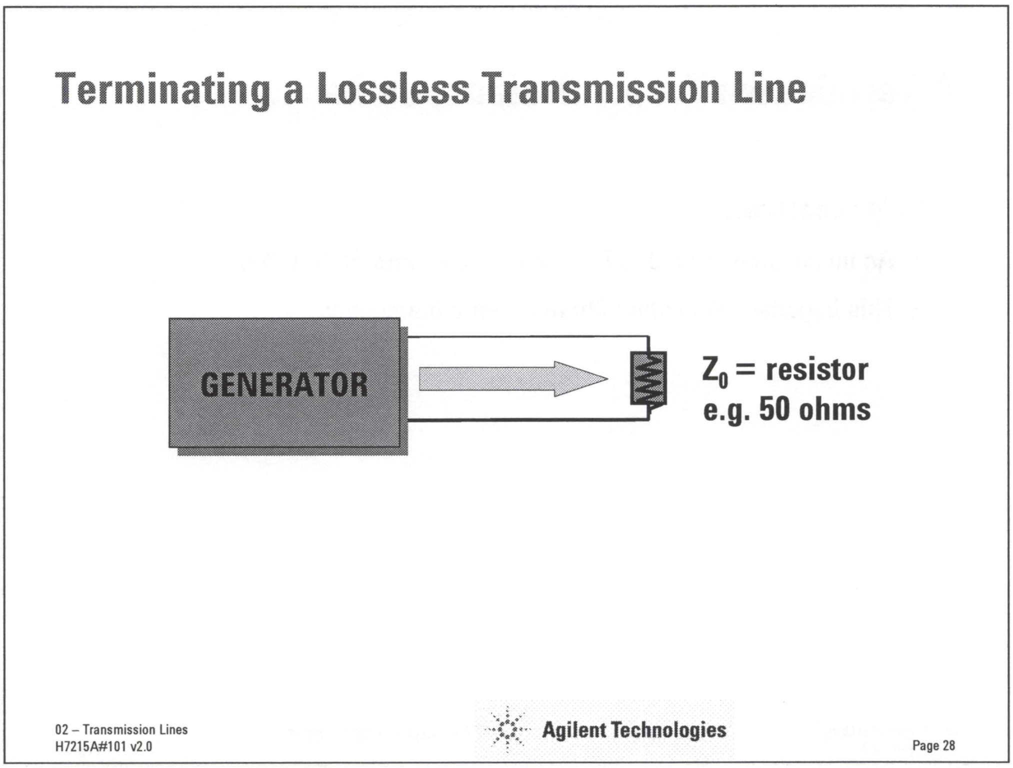
Page 27

It is important to remember that the characteristic impedance of a transmission line is related to its primary constants, and this impedance is a property of the line however long it is or however it is terminated. A normal impedance measurement of the line would render its characteristic impedance if the line were terminated in Z_0 or of infinite in length.

Looking ahead.

As we shall see the measured impedance at the input of a transmission line depends on characteristic impedance, load impedance and the length of the line.

We have only considered the incident wave so far, since only an incident wave can exist on Z_0 terminated line, this condition is sometimes called a *reflectionless termination*.



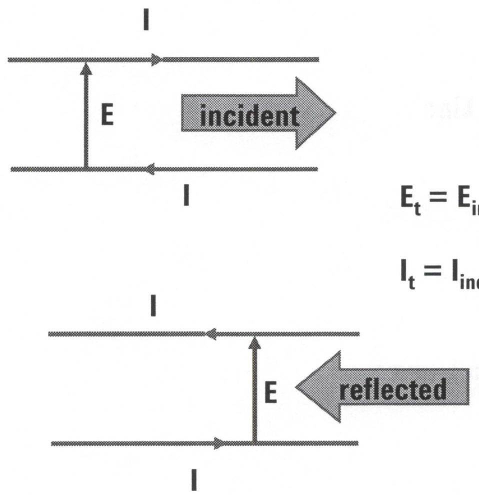
A very useful and important conclusion we made for a lossless line was that its characteristic impedance was real so a reflectionless termination for a "lossless" transmission line is a lumped resistive component.

In practice the creation of a suitable matched termination especially at very high frequencies is a bit more complicated. Lumped component resistors have more than just resistive properties at high frequencies.

Transmission Line Agenda

- Characteristic Impedance
- Terminating a Transmission Line
- Reflections
- Impedance
  - What is the Smith Chart ?
- Standing Waves
- Some practical Transmission Lines

Incident and Reflected Waves - How They Add



$E_t = E_{inc} + E_{ref}$

$I_t = I_{inc} - I_{ref}$

Incident Waves, Reflected Waves and Load Impedance

$$Z_L = \frac{V_L}{I_L} = \frac{V_i + V_r}{I_i - I_r}$$

substitute $$I_i = \frac{V_i}{Z_0}$$ and $$I_r = \frac{V_r}{Z_0}$$

02 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 31

The impedance at a point on a transmission line is defined as the ratio of the total voltage to the total current at that point on the line. If voltages add in the same sense then corresponding currents must subtract. For example when the current vectors add at a short circuit termination the voltage vectors subtract, and vice versa for an open circuit. It is logical therefore to write the above equation with opposite signs of V_r with respect to I_r.

The reflected wave is subject to the same characteristic impedance as the incident wave so we may substitute for I_i and I_r as shown.This will enable the expression of Z_L in terms of Z_0 and voltage ratios.

Complex value of Z_L

$$Z_L = Z_0 \frac{1 + \frac{V_r}{V_i}}{1 - \frac{V_r}{V_i}} = Z_0 \frac{1 + \Gamma}{1 - \Gamma}$$

02 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 32

Here we see that Z_L is a function of the complex ratio of Vr and Vi. We shall see the significance of this fact we discuss impedance vs. line length.

Reflection Coefficient and Load Impedance in Terms of Gamma

- The *reflection coefficient* is the ratio of the two vectors V_r and V_i and is given the symbol Γ (Gamma):

$$\Gamma = \frac{V_r}{V_i}$$

- Thus Z can be expressed in terms of Γ:

$$Z_L = Z_0 \frac{1 + \Gamma}{1 - \Gamma}$$

02 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 33

The impedance of a load termination is a function of reflection coefficient and characteristic impedance.

Gamma in Terms of Normalized Impedance

- **Normalized impedance**

$$\frac{Z_L}{Z_0} = Z_n$$

$$\Gamma = \frac{Z_n - 1}{Z_n + 1}$$

02 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 34

Gamma Γ may also be expressed in terms of the normalized impedance Z_n

$$Z_n = \frac{1+\Gamma}{1-\Gamma} \quad \text{if} \quad Z_n = \frac{Z_L}{Z_0}$$

This further simplification introduces another important concept:
Normalized Impedance Z_n.

Transmission Lines Agenda

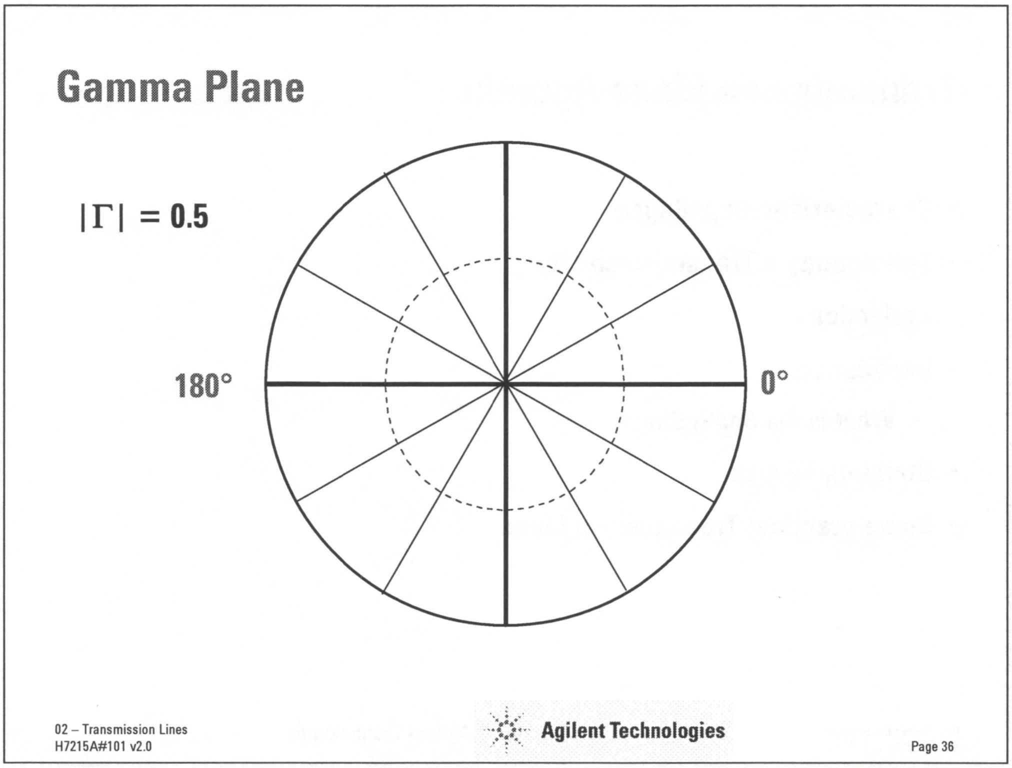
- Characteristic Impedance
- Terminating a Transmission Line
- Reflections
- Impedance
  - What is the Smith Chart
- Standing Waves
- Some practical Transmission Lines

02 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 35

No discussions of reflections and impedance at RF would be complete without some mention of the Smith Chart. The detailed use of the Smith Chart is outside the scope of this class but its usefulness should be appreciated.



Now we have an expression for the reflection coefficient in a polar coordinate form, the next step is to plot it on polar graph paper. This is the basis for a graphic calculator called the Smith Chart. Phillip Smith an engineer for the famous Bell Labs, developed his ideas over several years which resulted in an article about the chart first published in Electronics Magazine in 1939.

Lines of constant Rho

All lines of constant reflection coefficient are concentric circles. Shown is the $|\Gamma|$ = 0.5 line.

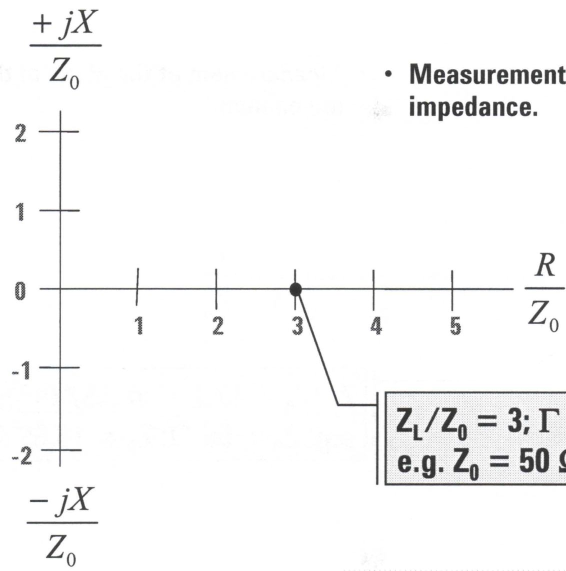
A variable resistance with no reactance.

For a purely resistive variable impedance load, imagine a variable resistance from infinite to zero ohms, i.e. from an open to a short, this would start at the right of the diagram and move along the horizontal axis to the left. As the locus coincided with the center, then the variable resistance would be at Z_0. This horizontal line is the *real* axis, notice the flip from 0° to 180° as the locus moves past the center

A variable reactance with no resistance.

If the locus of the load followed the outer circle then at all places there is 100% reflection, ($|\Gamma|$ = 1), but only on the horizontal axis is it resistive, in the upper part of the circle the reflection would be purely inductive and on the lower part purely capacitive.

Normalized Impedance Plane 1



- **Measurement at the plane of the impedance.**

$Z_L/Z_0 = 3$; $\Gamma = 0.5\underline{/0^\circ}$;
e.g. $Z_0 = 50\ \Omega$; $Z_L = 150\ \Omega$

02 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

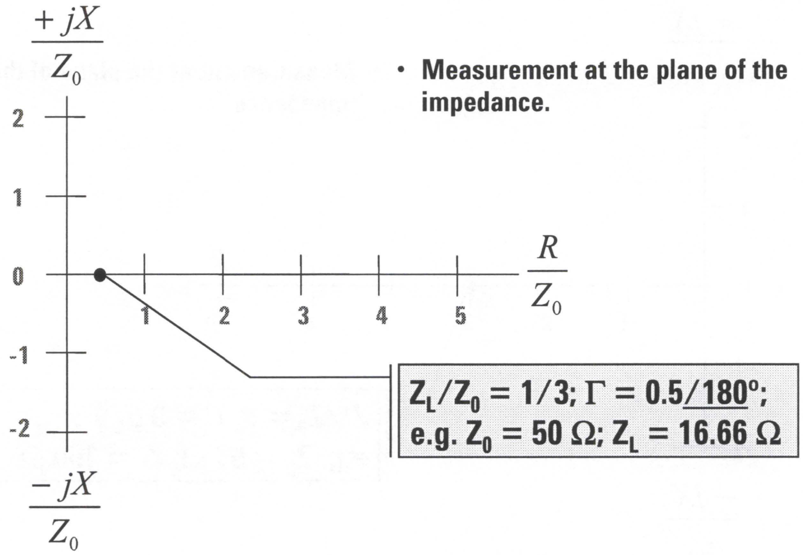
Page 37

To understand the great importance of the Smith chart, first look at a more traditional graph; the rectangular impedance graph.

Shown is the point for a normalized real impedance of 3. This corresponds to a resistance of three times the characteristic impedance. The reflection coefficient is +0.5 which is $0.5\underline{/0^\circ}$

Normalized Impedance Plane 2

- **Measurement at the plane of the impedance.**



02 – Transmission Lines
H7215A#101 v2.0

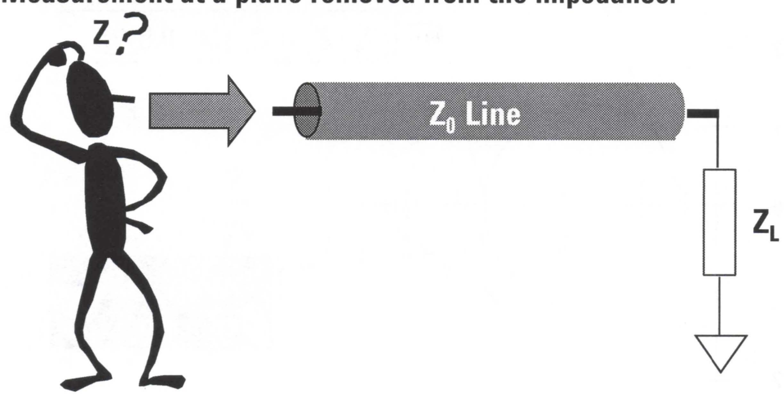
Agilent Technologies

Page 38

This time an normalized impedance of 1/3 is plotted on the same chart. The calculated reflection coefficient is now –0.5 which is $0.5\underline{/180^\circ}$ expressed in polar coordinates.

Impedance and Line Length

- **Measurement at a plane removed from the impedance.**

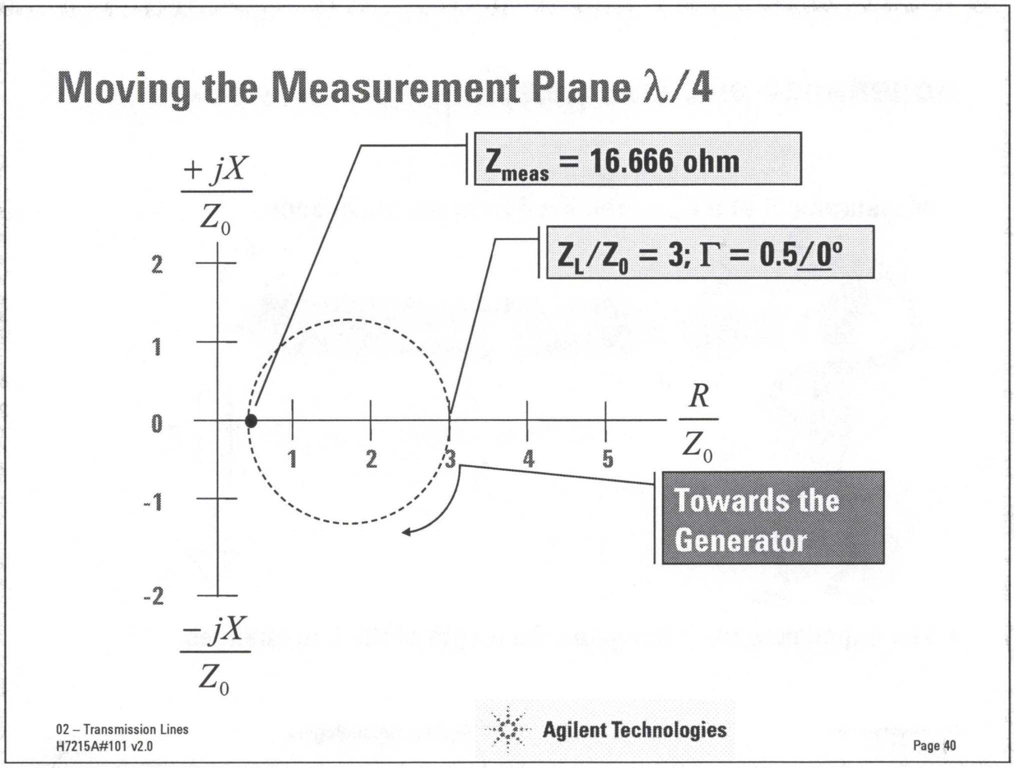


- **The impedance will change as the length of the line changes.**

In the case for the previous example. The only difference was the angle between the reflected and incident waves at the measurement point, ρ(rho) is constant.

We could have measured at the plane of the termination and changed the value of the termination from 3 to 1/3 Z_n. Or ...

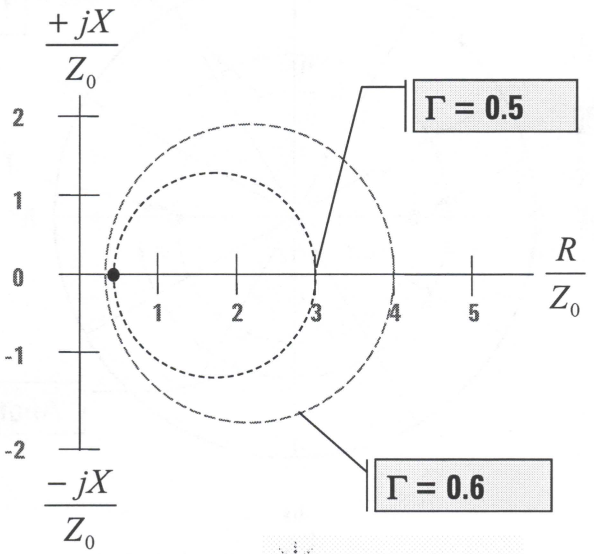
Have a constant normalized impedance 3, at the termination and move the measurement point by 1/4 wavelength toward the generator. This would change the total phase angle between the incident and reflected wave by 2 × the physical angular distance. i.e. 1/2 wavelength or 180°.



Since the reflection coefficient has the same magnitude in both cases, it can be shown, but not here, that as the measurement point is moved away from the termination towards the generator the *locus* of the impedance will follow a *circle* as shown.It becomes capacitive, as it cuts the real axis again it is resistive Zn = 1/3 +j0. Then with further movement it becomes inductive before it once again becomes real at Zn = 3 + j0.

This chart has a problem. It would be very impractical to plot very high impedances on this chart. The previous example of varying a resistance from open to short could not be done. Also other Gammas will produces other circles, but they are non-concentric.

Impedance Plane with Different Reflection Coefficient Circles



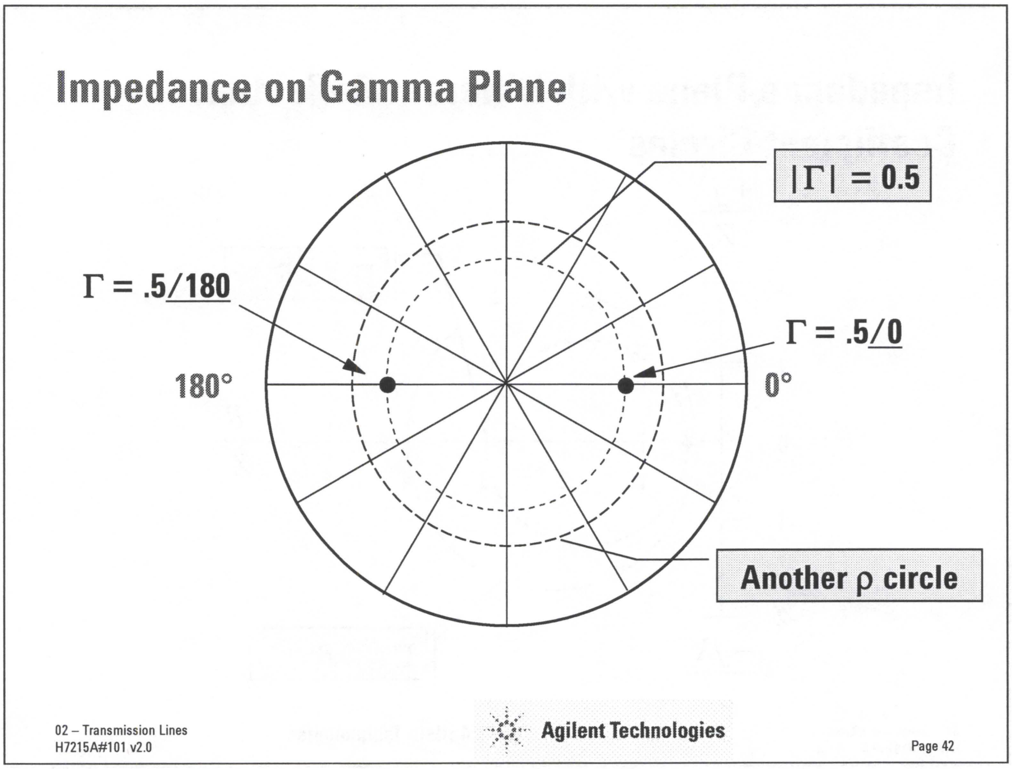
02 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 41

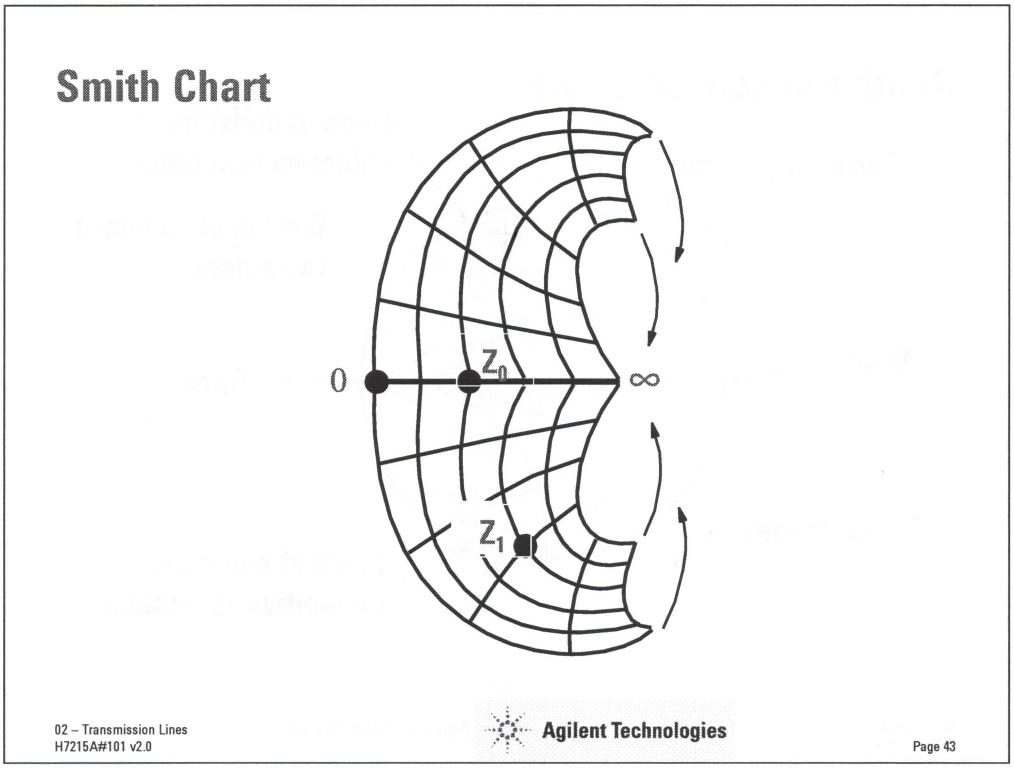
Since the reflection coefficient has the same magnitude in both cases, it can be shown, but not here, that as the measurement point is moved away from the termination towards the generator the *locus* of the impedance will follow a *circle* as shown.It becomes capacitive, as it cuts the real axis again it is resistive Zn = 1/3 +j0. Then with further movement it becomes inductive before it once again becomes real at Zn = 3 + j0.

This chart has a problem. It would be very impractical to plot very high impedances on this chart. The previous example of varying a resistance from open to short could not be done. Also other Gammas will produces other circles, but they are non-concentric.

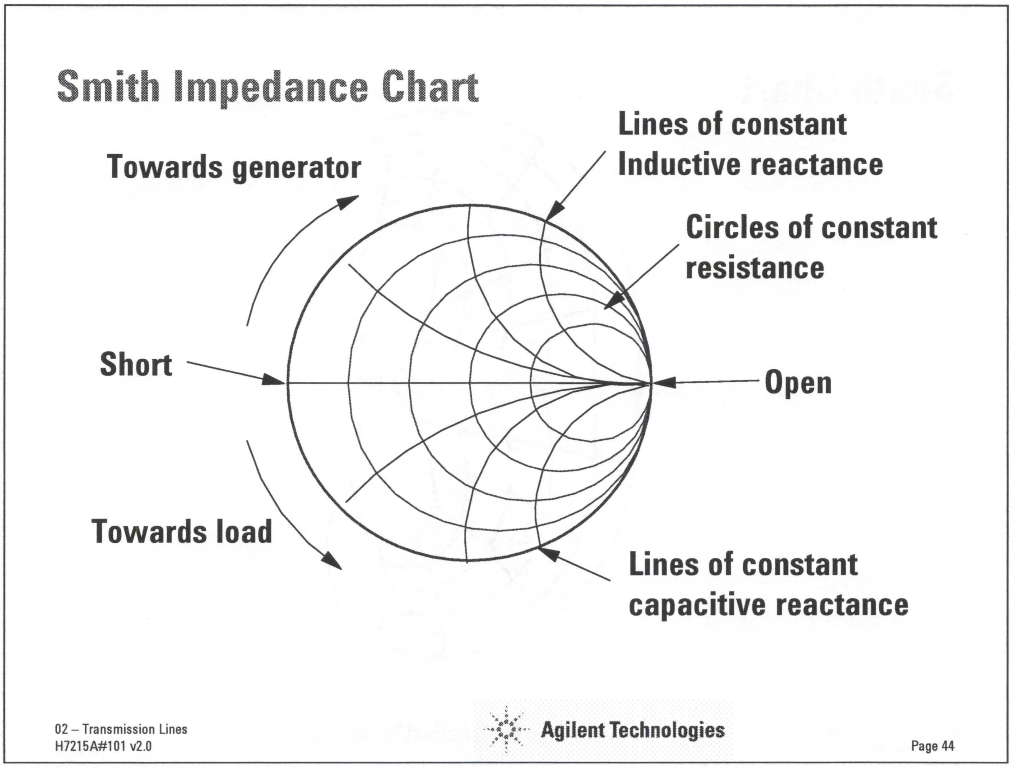


On the polar plot the locus of Gamma due to changing the measurement plane with respect to the termination also follows a circle. Any different Gammas will produce concentric circles. Corresponding impedances can be plotted whatever their values.

The other scalar reflection measures, SWR or Return Loss may be represented as a set of concentric circles.



The Smith chart is a transformation of the impedance chart, imagine it stretched around to force it into the polar Gamma plane.

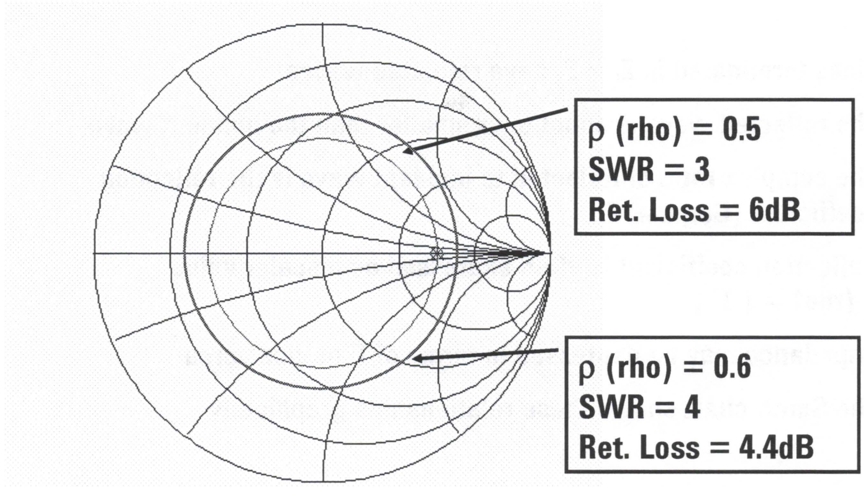


This is the layout of the Smith Chart , it has the best features of a polar plot with the impedance chart and can enable RF engineers to move from reflection measurement to impedance evaluation without complicated calculations. The chart will easily handle both large and small impedances on the same chart.

There are many variations of the Smith Chart, for example the mirror image f the above becomes the Smith admittance chart.

Many RF and microwave design books will explain this most useful tool. A source book is one by Phillip Smith himself "Electronic Applications of the Smith Chart" Noble Publishing, Atlanta, GA.

Smith Chart Example

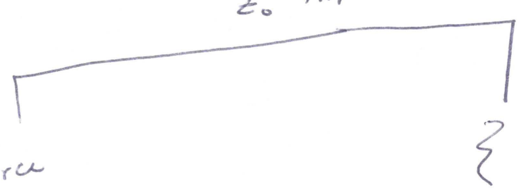


02 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 45

Z_0 Impedance



(Volt) E Field returns if open

(Current) I Field returns if short ~~180° out of phase~~

Source

open/short

Returns 180° out of phase at short/open

$RL = -20 \log \rho$

Impedances and Reflections

- Lines terminated in $Z_L \neq Z_0$ have reflected waves
- The reflected wave is equal to or smaller than the incident wave.
- The complex ratio of reflected to incident wave is the reflection coefficient (Gamma) Γ
- Reflection coefficient is often expressed as a scalar value ρ (rho) = $|\ \Gamma\ |$
- Impedance may be expressed in terms of Γ or vice versa
- The Smith chart shows these relationships graphically

02 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 46

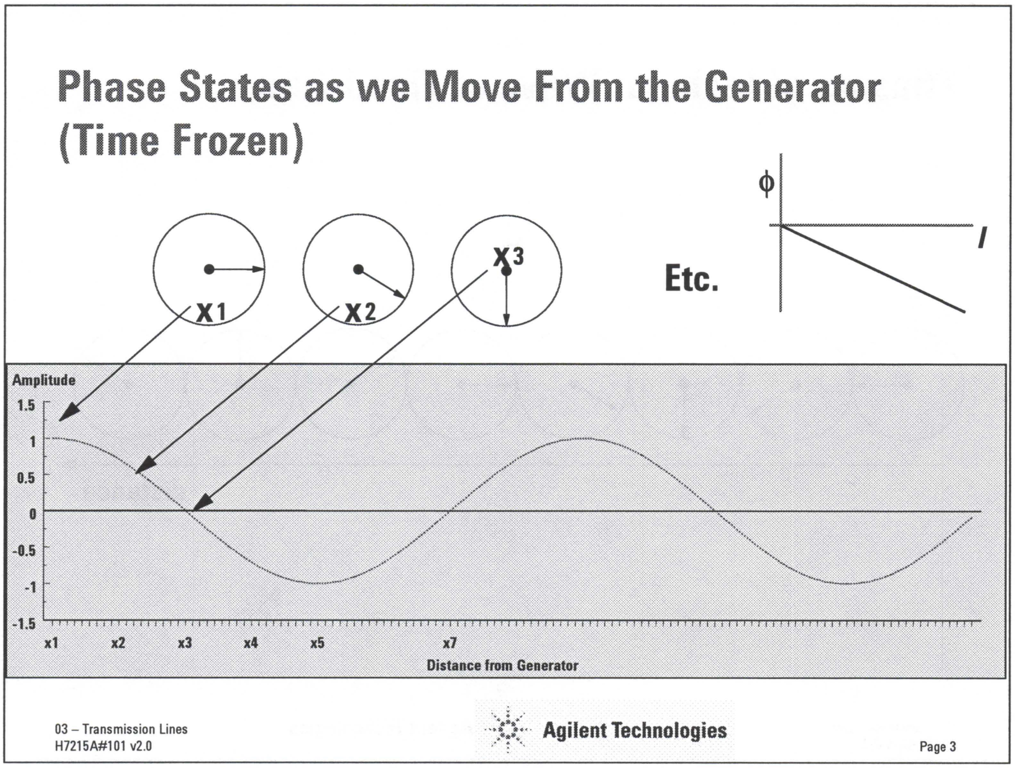
RF & Microwave Measurement Fundamentals

3 — Transmission Lines

Agilent Technologies

Transmission Line Agenda

- Characteristic Impedance
- Terminating a Transmission Line
- Reflections
- Impedance
- What is the Smith Chart
- Standing Waves
- Some practical Transmission Lines

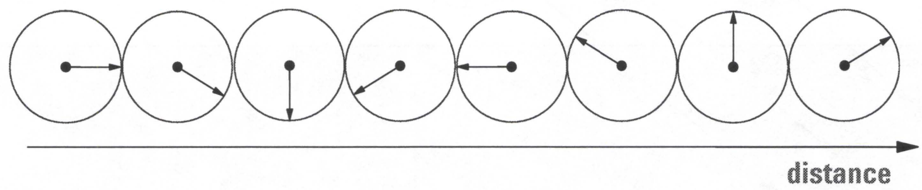


Before trying to explain the way in which the incident and reflected waves combine it is best to spend some time to understand the vector (phasor) diagram. This is the best way to show the progress of a wave in both distance and time and is also a very powerful model to apply to other aspects of RF technology.

The phasor, by definition, rotates in a *counter clockwise* direction with time. A constant frequency, constant amplitude sinewave, $f(t) = A\cos(\omega t)$, is imagined as a vector, length A, spinning around a fixed point with a rotational speed of ω radians/sec. When ωt is 0 rads. i.e. when $t = 0$ the vector is defined by a line from 0 (the axis) to A on the horizontal axis.

The drawing above however shows *distance* not time, when the distance from the generator is such that the angle is $-\pi/2$ this must represent the state of the wave before $t = 0$. (If time is "frozen" then the state of the vector down the line represents the "history" of the wave as it was at the generator end of the line.) this is why the successive stages are shown as a *clockwise* progression.

Phasors of Incident Wave at Fixed Time

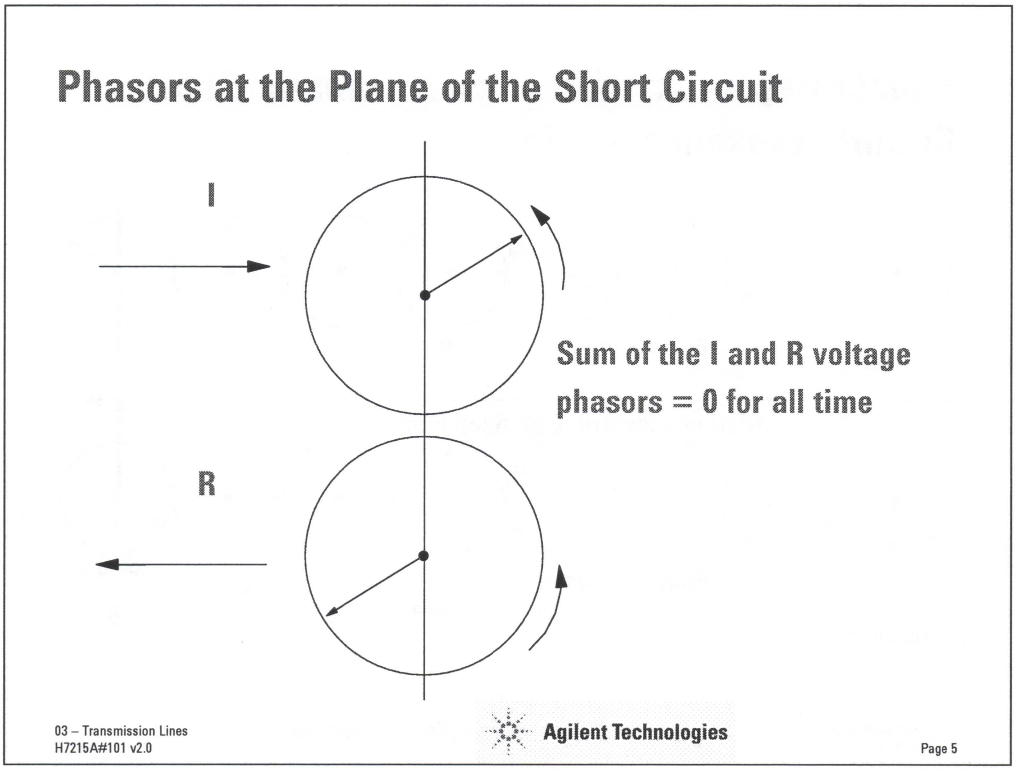


03 – Transmission Lines
H7215A#101 v2.0

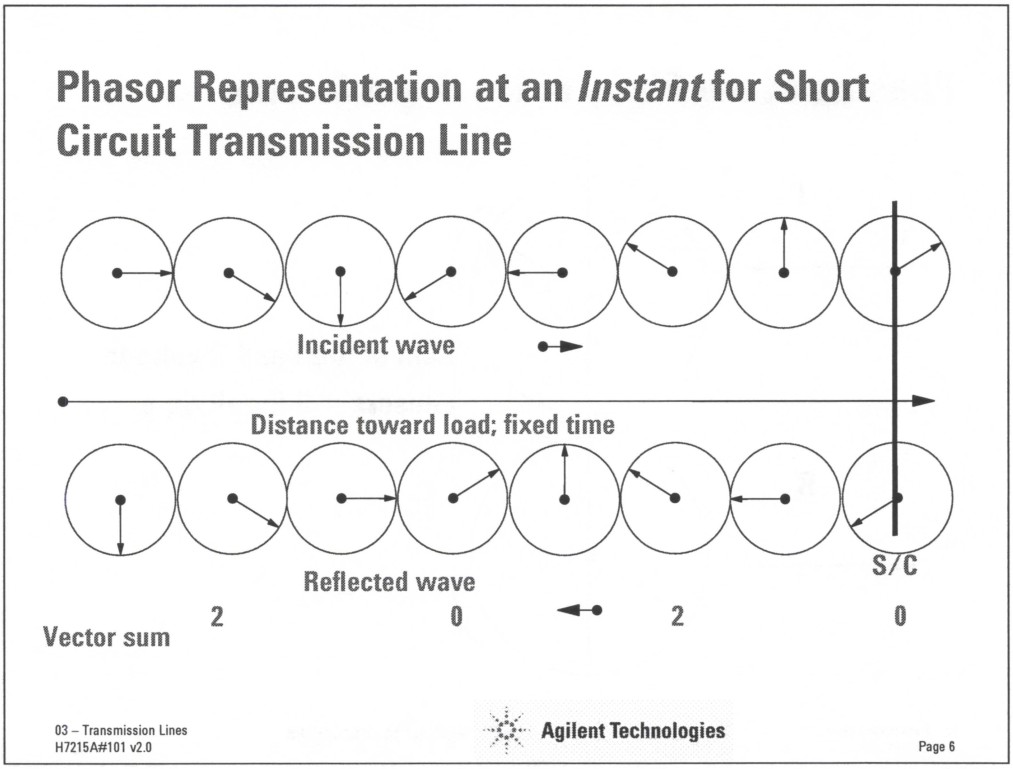
Agilent Technologies

Page 4

This diagram can be repeated at various distances down the line.



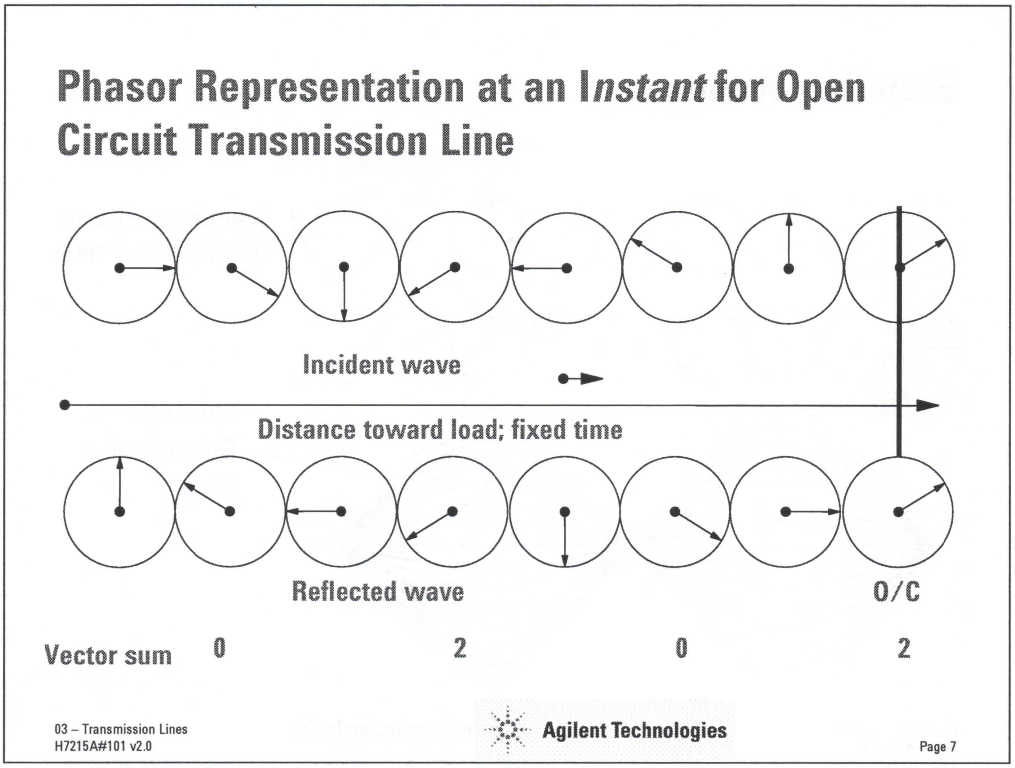
If the wave meets a short circuit then there is a phase reversal, at the plane of the short circuit.



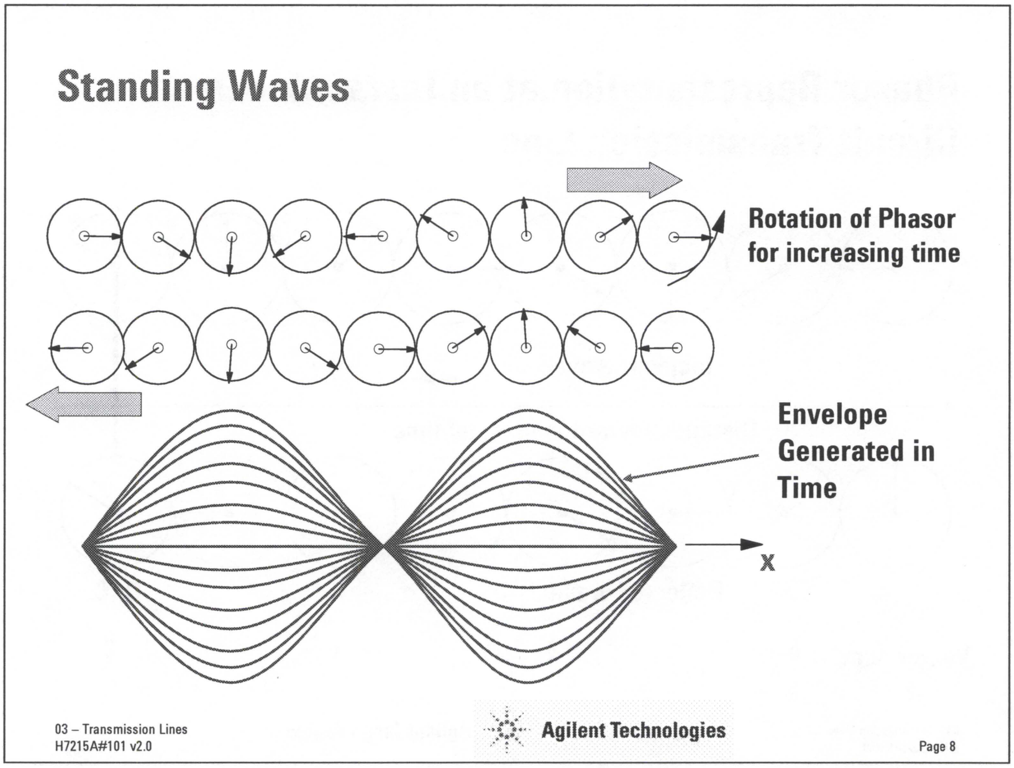
The total voltage at a point on the transmission line is the vector sum of all the waves present at that point. In this case this is the sum of the incident and reflected waves.

Imagine all the vectors shown rotating as time progresses (ccw). It can be seen that the vector sum of the incident and reflected waves form a pattern that is static with time.

The vector magnitudes at each point are constant, at some points the sum is zero and at other points the sum is two, between those points there is are intermediate values.

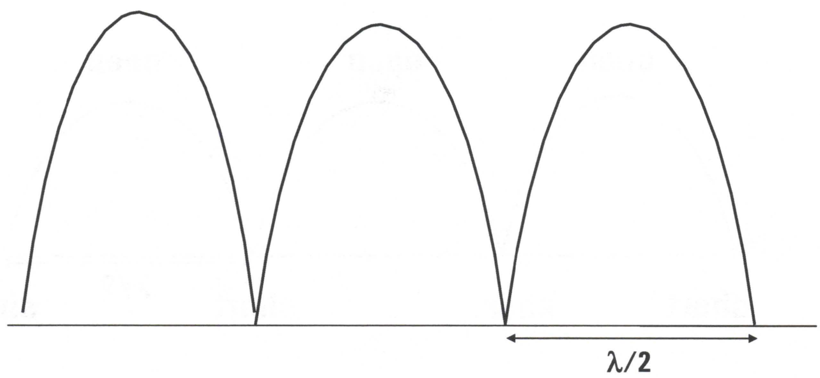


With the open circuit there is an identical pattern, The difference that the vector sum at the plane of the open circuit is two.



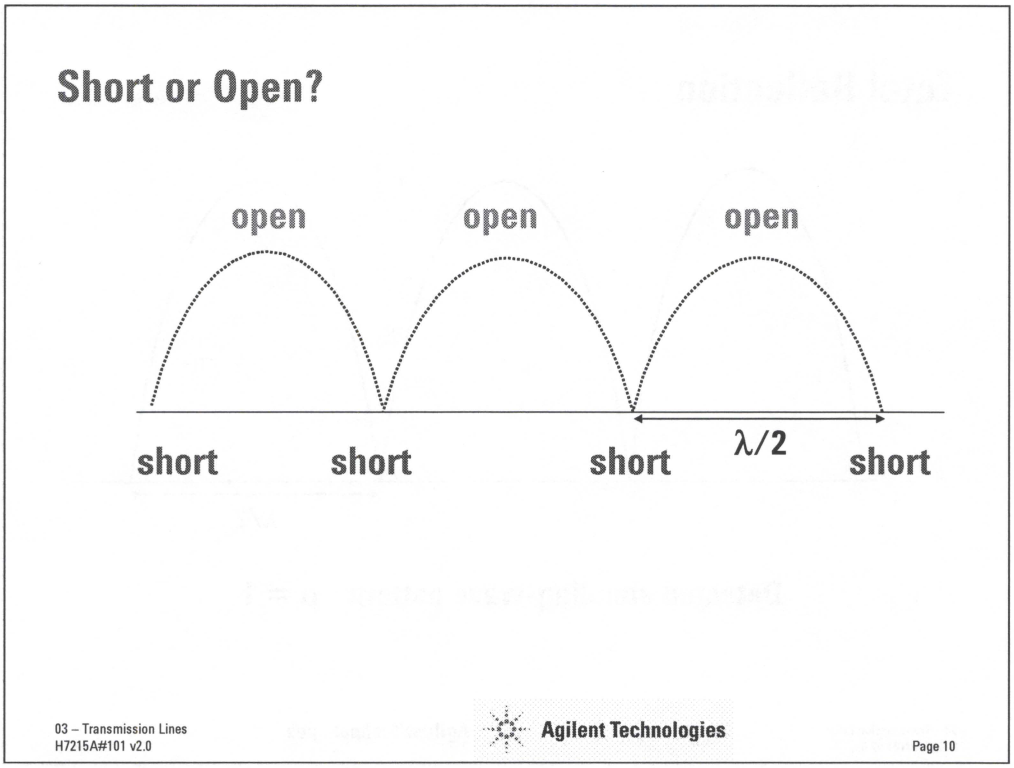
When evaluating the *amplitude* of a phasor we must remember that this will vary f(t) = A.cos(ωt), as we explore the vector sum of the waves it is the "A" that varies between zero and two. This phenomena is called a ***standing wave***.

Total Reflection



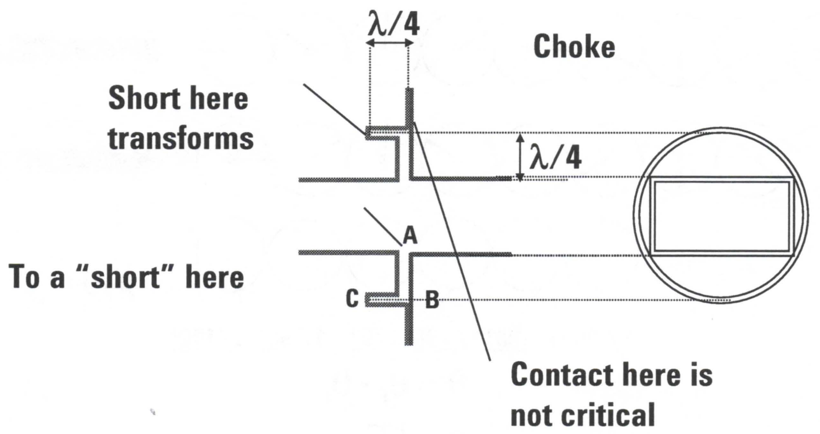
Detected standing-wave pattern $\rho = 1$

The standing wave is an alternating signal. The magnitude of the standing wave as a function of position is of interest. The shown standing wave pattern is for a total reflection. The minimum positions are called *nodes* and the maximums *antinodes.* The distance between adjacent nodes or antinodes is one half wavelength.



The cyclical nature of the pattern shows how the perceived impedance changes as the observation is made in different positions along the line. In the extreme a short is an open one quarter wavelength away. This is of practical use in waveguide where an open is needed for network analyzer calibration, but a waveguide open is not realizable. So a so called offset short is used instead, this would be a short a quarter wavelength from the plane of the required open circuit.

Waveguide Choke Coupling



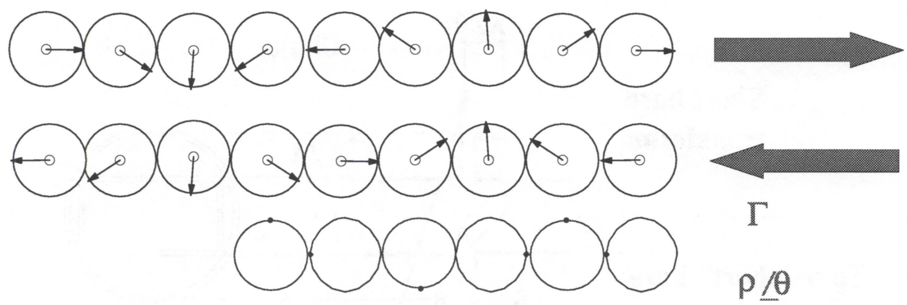
03 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 11

The waveguide choke coupling is a good example of using the $\lambda/2$, $\lambda/4$ transformation to practical advantage. Because there will be some uncertainty of contact between the waveguide flanges, transforming the physical short at C by $\lambda/2$ to a virtual short at A, while the contact at B is not so critical because the transformation by l/4 means there is a virtual open circuit at B thus wall currents are small (zero) at the point of contact.

Total Reflection with Gamma



1/ 90° 1/180° 1/ -90° 1/0° 1/ 90° 1/180°

$$\theta = \theta_r - \theta_i$$

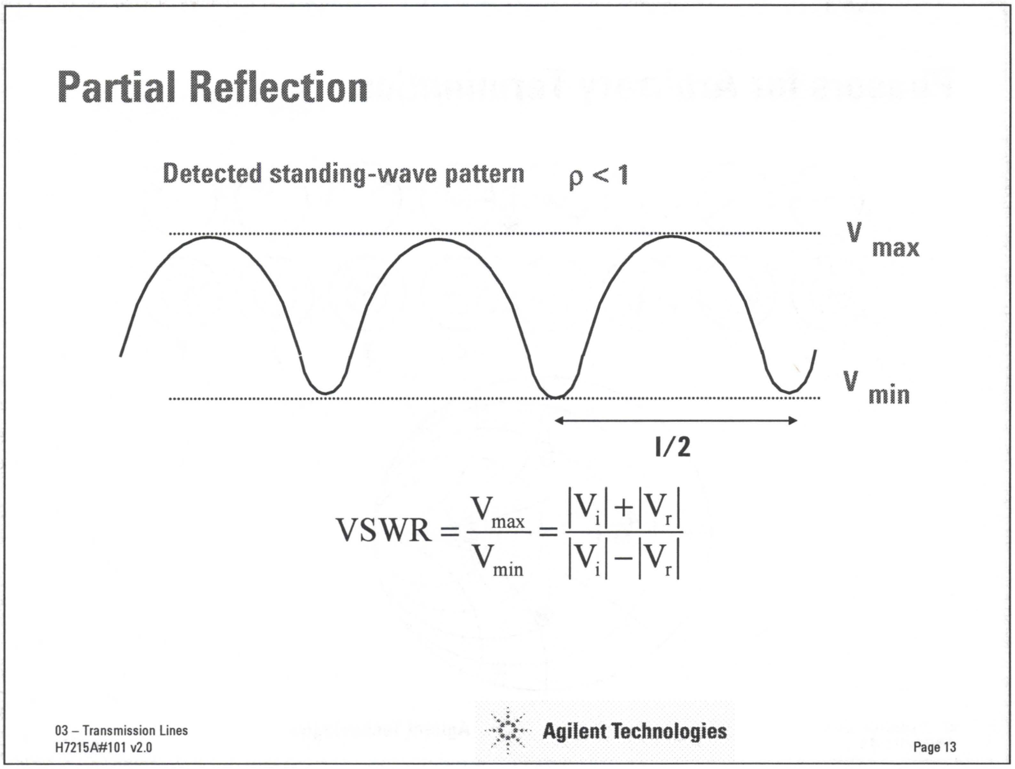
$$\rho = |\Gamma|$$

03 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 12

In this diagram the value for gamma is plotted on the polar(Smith) chart. Note the magnitude of the reflection coefficient is constant, but the angle varies.



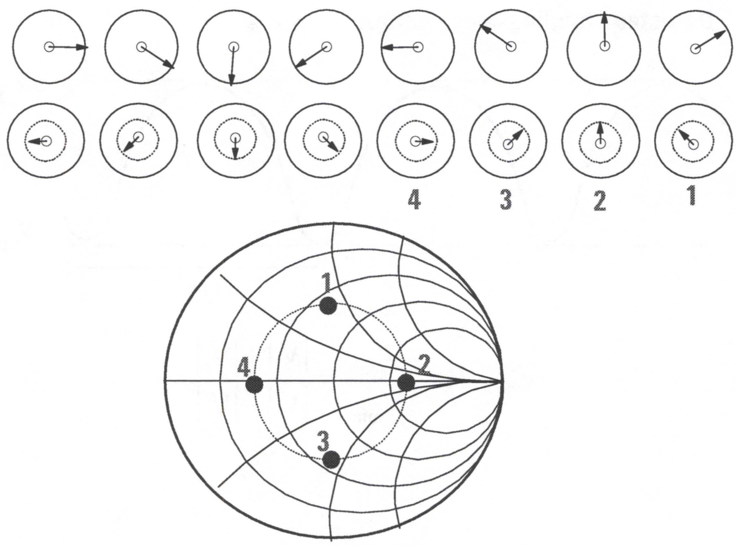
When the reflection is not 100% there are still standing waves, this time the nodes do not reduce to zero.

This is a good time to introduce the definition of a quantity known as VSWR, or just SWR and sometimes called "swirr."

SWR is the ratio of the voltage at the antinode to the voltage at the node.Its relationship to the magnitudes of the incident and reflected waves is simple:

$$VSWR = \frac{V_{max}}{V_{min}} = \frac{|V_i| + |V_r|}{|V_i| - |V_r|}$$

Phasors for Arbitrary Termination



03 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 14

For an arbitrary termination, the magnitude of the reflected wave is smaller than that of the incident wave. The magnitude of the reflection coefficient is therefore less than one. The impedance varies as the reflection is viewed from different positions along the line.

The Slotted Line

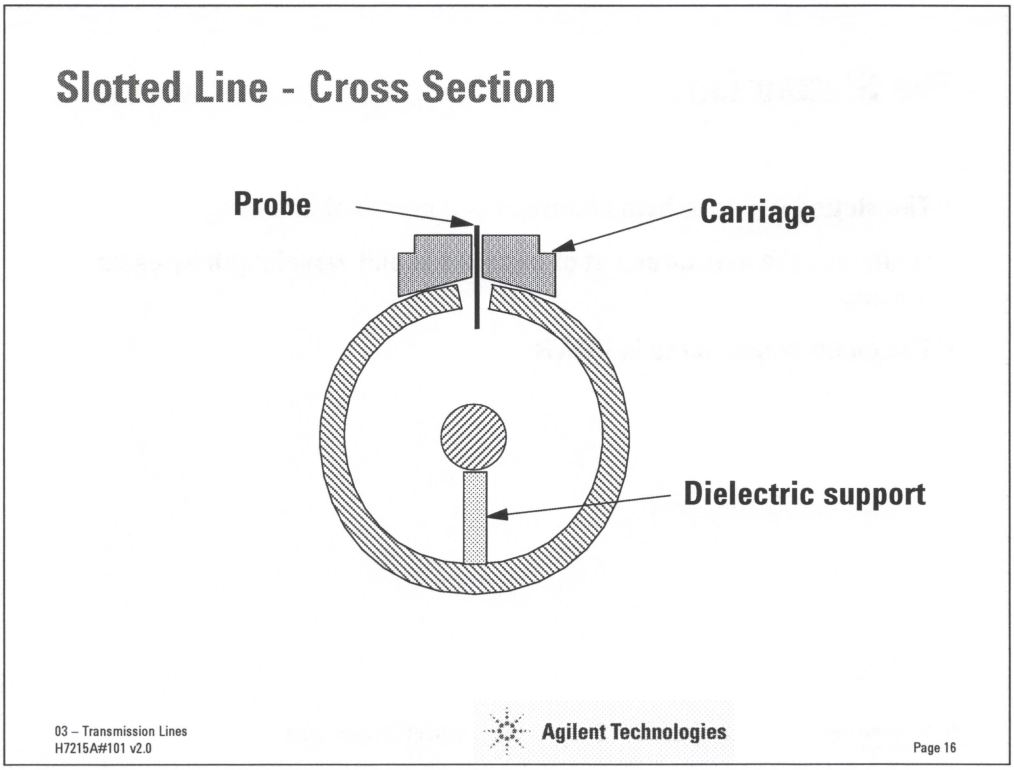
- **The slotted line is of both historical and practical interest.**
- **It allowed the measurement of impedance and wavelength by basic means.**
- **The quantity measured is VSWR.**

03 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 15

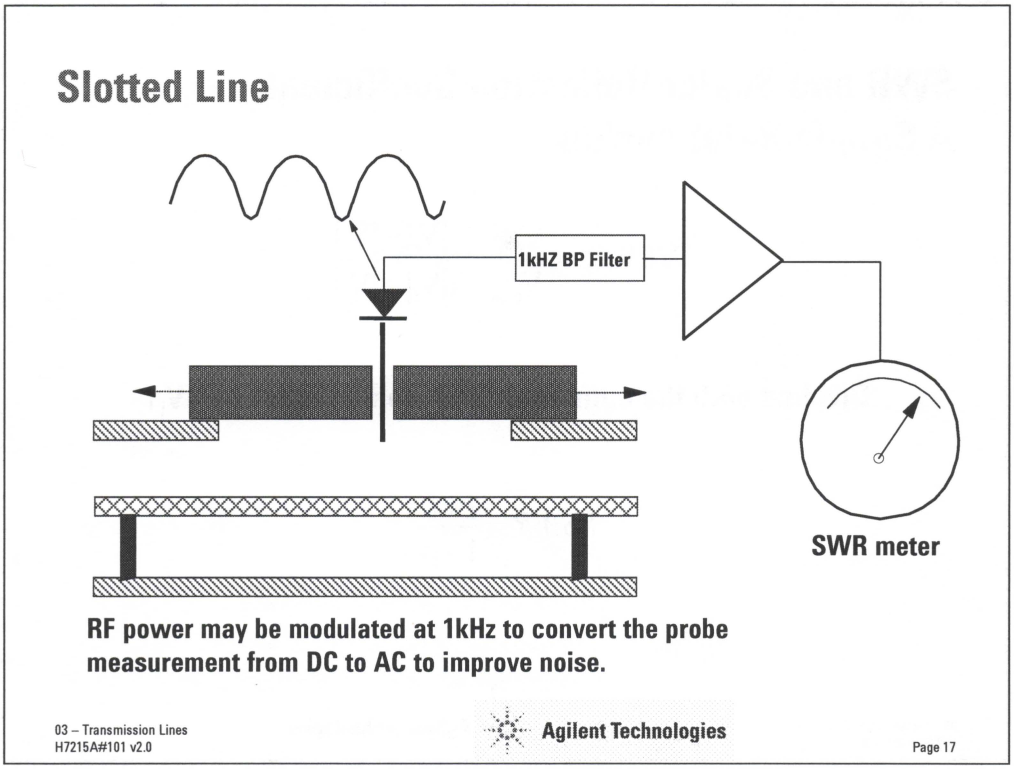
A small digression into some historical background. The term VSWR is still used in the microwave and RF business, although it is very rarely measured directly, when VSWR is measured directly the device used is called slotted line.



This example shows a cross-section of a special coaxial section of line. It is mainly an air filled line with small support rods or discs for the center conductor.

A narrow slot is cut into the outer shield in the longitudinal direction. Since in the fundamental mode of propagation all currents flow longitudinally, this slot has a minimal effect on current flow.

Through the slot a small probe is inserted to sample the RF field , this probe must be small so that there is no disruption to wave propagation.



This longitudinal cross-section of the slotted line shows how the standing wave is measured. As the carriage moves the meter will record the maximum and minimum values of the standing wave.

SWR and Scalar Reflection Coefficient: A Simple Relationship

$$VSWR = \frac{V_{max}}{V_{min}} = \frac{|V_i| + |V_r|}{|V_i| - |V_r|}$$

Dividing both the numerator and denominator by $|V_i|$

$$VSWR = \frac{1+\rho}{1-\rho}$$

03 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 18

Standing wave ratio and reflection coefficient have a simple relationship shown by the above equation. But the range of values is very different. If ρ is 1 (100% reflection) then VSWR is infinite. If r is zero (perfect match) then VSWR is 1. Often SWR is written with the ratio symbol (:) such as 1.5:1 or 2:1. Because SWR is still in frequent use it is important to be able to easily convert from SWR to scalar reflection coefficient ρ and vice versa.

$\rho = \frac{V_R}{V_I}$

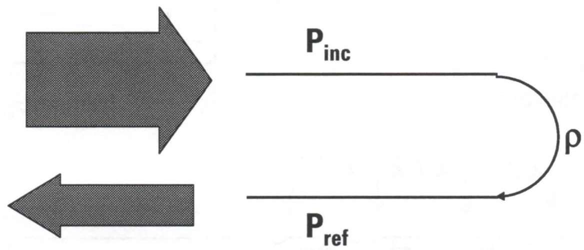
$RL = -20 \log \rho$

$\rho = |\Gamma|$

$VSWR = \frac{1+\rho}{1-\rho}$

Return Loss

$$\text{Return Loss} = -20 \cdot \log(\rho)\ \text{dB} = -10 \cdot \log(P_{ref}/P_{inc})$$



03 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 19

Return loss has a physical meaning, it is how much POWER is lost between the incident and reflected wave.

- e.g.. a 3dB return loss means that the reflected power is half the incident.

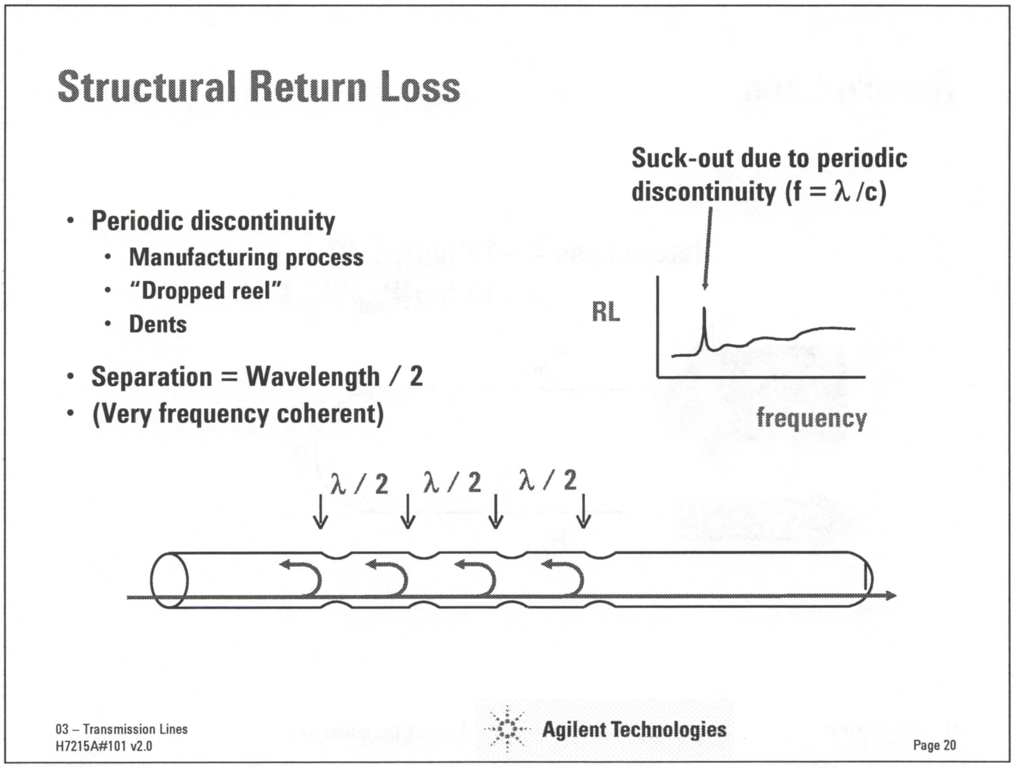
The minus sign ensures that the expression of return loss will always be positive, since $\rho<1$ means $\log(\rho)$ is negative thus the expression $-20\log(\rho)$ is positive.

A "good match" i.e. low reflection has a high return loss, e.g.. a good termination may have a return loss of 30dB.

- Total reflection;   RL = 0
- Perfect termination RL = infinity

Return Loss may also be encountered at low frequencies as a way of expressing impedance relative to a reference. If this reference impedance is Z_0 then the previous definition holds.

$$\text{Return Loss} = 20 \log \left| \frac{Z_1 + Z_2}{Z_1 - Z_2} \right|$$



Structural Return Loss and long cables

Structural return loss is a special case of return loss applied to cables that are many wavelengths long. Physical deformation of the cable, by handling or manufacturing process, cause reflections. Structural return loss shows at particular frequencies where the total reflected wave is the sum of the individual reflections from periodic deformations occurring at a half-wavelength spacing.

A narrow spike of high return loss results from a specific frequency component being reflected from the cable and not transmitted through the cable. When the effects from these small discontinuities are summed, they represent a major effect on the transmission capability of the cable. Periodic spacing of irregularities in time or distance leads to a cumulative effect in frequency.

This type of return loss has been observed also by workers with semi-rigid cable. If equal length semi-rigid units are cascaded the connections will occur at regular distances and cause SRL.

Structural Return Loss Definition

- **Measure impedance versus frequency; Z_0 is system impedance**

$$Z_{in}(\omega) = Z_0 * \frac{(1+\rho(\omega))}{(1-\rho(\omega))}$$

- **Compute average cable impedance**

$$Z_{cable} = \frac{\sum |Z_{in}(\omega)|}{N}$$

- **Compute reflection coefficient relative to average cable impedance**

$$\rho_{SRL}(\omega) = \frac{Z_{in}(\omega) - Z_{cable}}{Z_{in}(\omega) + Z_{cable}}$$

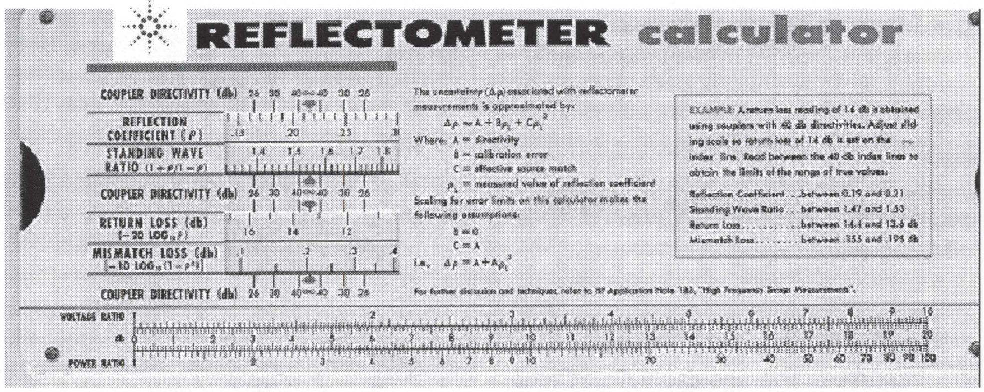
03 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 21

The definition of structural return loss is a variation of the return loss definition, but is based on the cable's actual (average) impedance which may not be exactly Z_0.

The Reflectometer Calculator



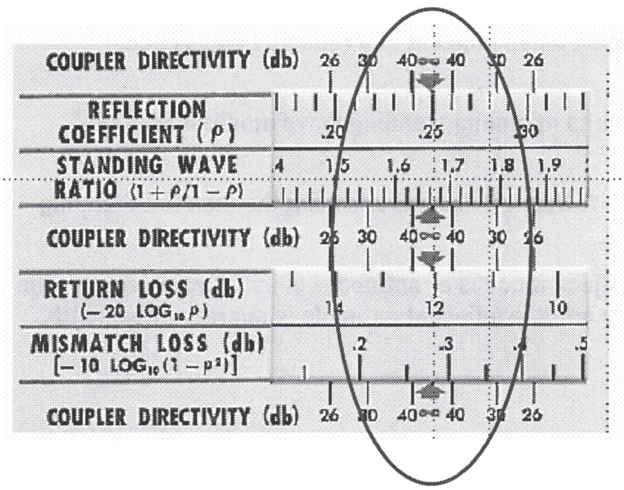
03 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 22

This tool is the reflectometer calculator, it has many uses but one we will explore now is to convert between VSWR, scalar reflection coefficient and return loss.

Converting between Rho, Return Loss and SWR



- For example:
  - $\rho = .25$
  - **SWR = 1.66**
  - **R-Loss = 12dB**

03 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 23

About Standing Waves

- Standing waves are formed when incident and reflected waves occur together.
- The ratio of the maximum to minimum standing wave amplitudes is the Standing Wave Ratio.
- If the line is long enough to support several wavelengths then the Standing wave pattern will repeat.
- The distance between adjacent nodes or antinodes is half a wavelength. The distance between a node and the adjacent antinode is quarter wavelength.

Effect of movement from the plane of the load

- Along a lossless line
  - Incident wave amplitude is constant
  - Reflected wave amplitude is constant
  - Line characteristic impedance (Z0) is constant
  - Phase difference between Ei and Er varies
  - Γ and Z measured away from the plane of the load varies

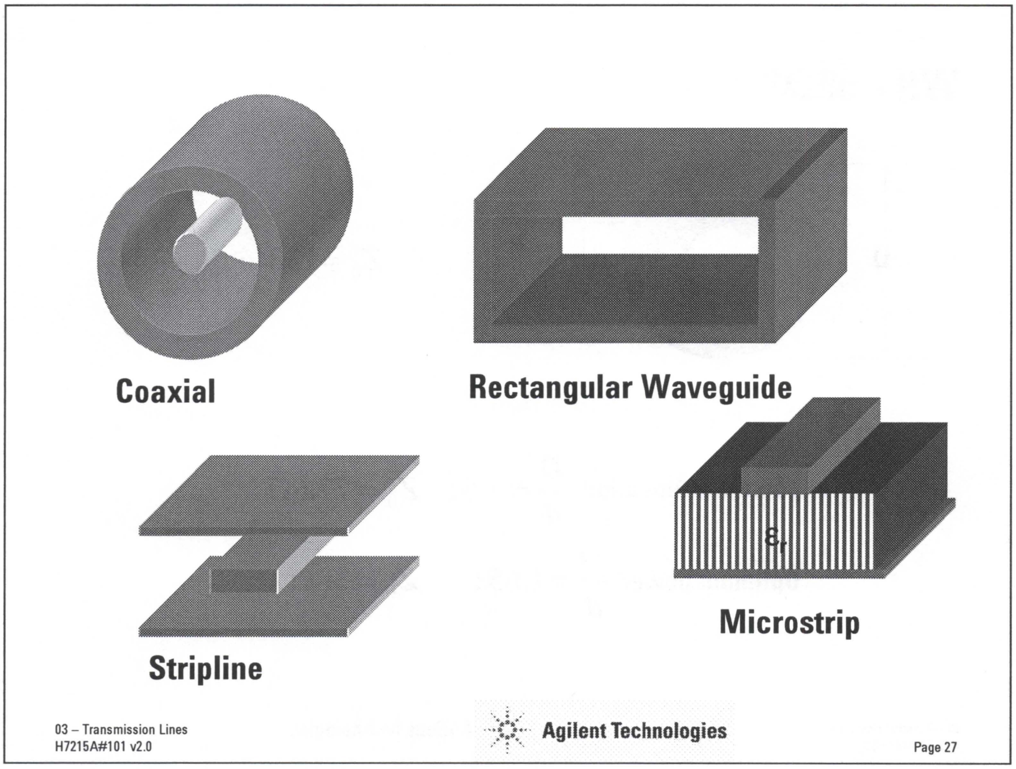
Transmission lines

- Characteristic Impedance
- Terminating a Transmission Line
- Reflections
- Impedance
- What is the Smith Chart
- Standing Waves
- Some practical Transmission Lines

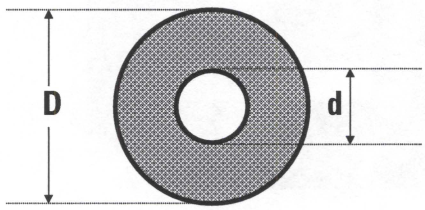
03 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 26


Coaxial
Rectangular Waveguide
Stripline
Microstrip
ε_r
03 – Transmission Lines
H7215A#101 v2.0
Agilent Technologies
Page 27

WHY 50Ω?



$$Z_0 = 138 \log_{10}\left(\frac{D}{d}\right)$$

Lowest attenuation $\frac{D}{d} = 3.6;\quad Z_0 = 77.6\Omega$

Optimum power $\frac{D}{d} = 1.65;\quad Z_0 = 30\Omega$

03 – Transmission Lines
H7215A#101 v2.0

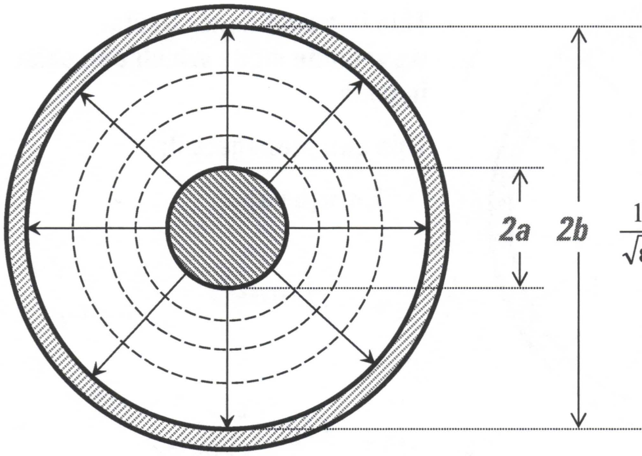
Agilent Technologies

Page 28

The range of impedance for reasonable mechanical ratios, for example if D/d varies from 1.5 to 10 the impedance for an air dielectric varies from 24.3 ohm to 138 ohms. So any mechanically easily made coax line is going to have a characteristic impedance around 25 to 100 ohms. For 50 ohms the ratio is 2.3.

This has not answered the question; why 50ohms? It seems to be a compromise.The lowest attenuation occurs when the inner to outer diameter ratio is 3.6 which corresponds to an impedance of 77.5 ohms. On the other hand, because the electric field strength is highest near the surface of the inner conductor this field strength could cause breakdown if too high. By calculating between the voltage breakdown in air and the transmitted power in a reflectionless line gives a diameter ratio of 1.65, corresponding to 30 ohms.

Coaxial TEM Mode



$$Z_0 = \frac{138}{\sqrt{\varepsilon}} \log_{10} \frac{b}{a}$$

$$\frac{1}{\sqrt{\varepsilon}} = \textit{velocity factor}$$

$$= \textit{1 for vacuum (air)}$$

$$\approx \textit{0.65 for teflon (PTFE)}$$

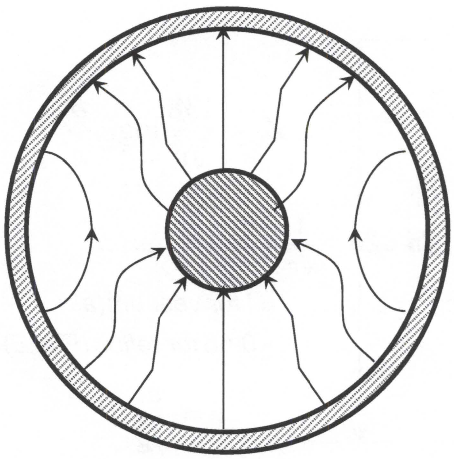
$$v = \frac{c}{\sqrt{\varepsilon}}$$

03 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 29

Coaxial TE_{11} Mode



- **This is a dominant circular waveguide mode which can exist in coax.**
- **This can propagate if:**
  - **$\lambda < \pi(b+a)$**

b = diameter inner circle
a = " outer circle

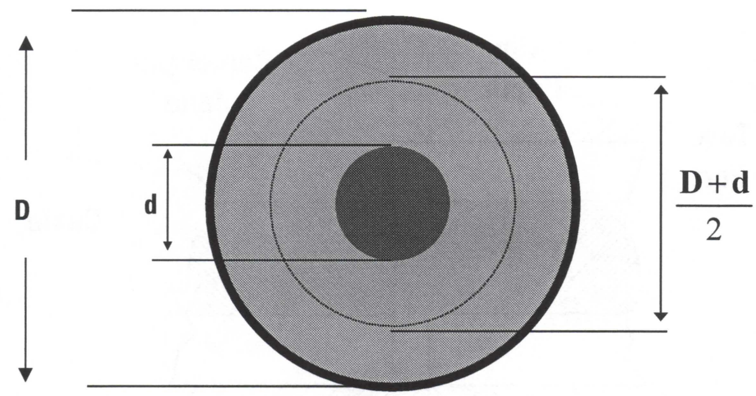
why HKL FRQ DON'T WORK IN COAX

03 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

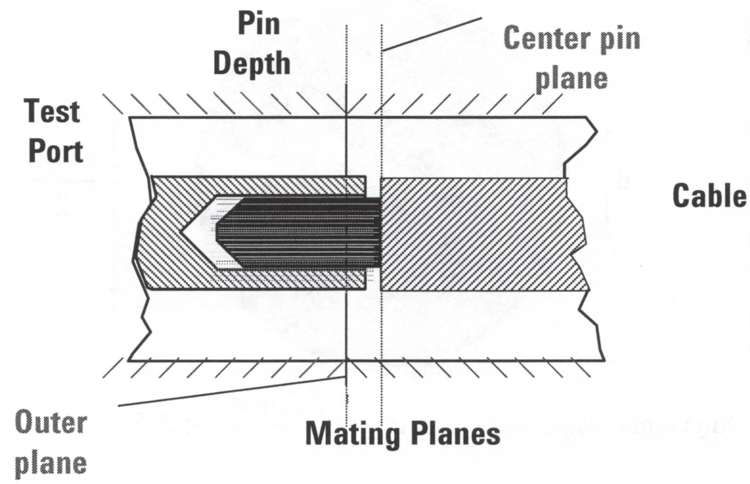
Page 30

Frequency Limit of Coax



High order modes may occur if: $\pi\left(\dfrac{D+d}{2}\right) \geq \lambda$

Connector Mating Planes



03 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 32

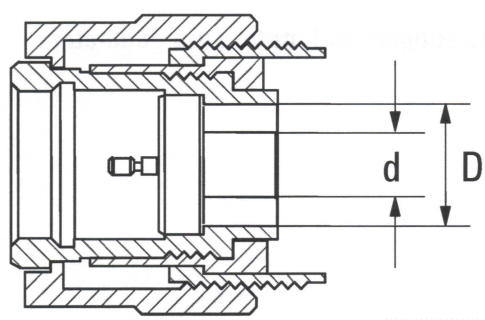
Characteristic Impedance

- **Model for Characteristic Impedance, Z (Low-Loss Case)**

$$Z_0 = \frac{1}{2\pi}\sqrt{\frac{\mu}{e}}\,\ell n\left(\frac{D}{d}\right)$$

D = Inner diameter of outer conductor

d = Outer diameter of inner conductor



D = 7.0 mm
d = 3.04 mm
Zo = 50 ohms

03 – Transmission Lines
H7215A#101 v2.0

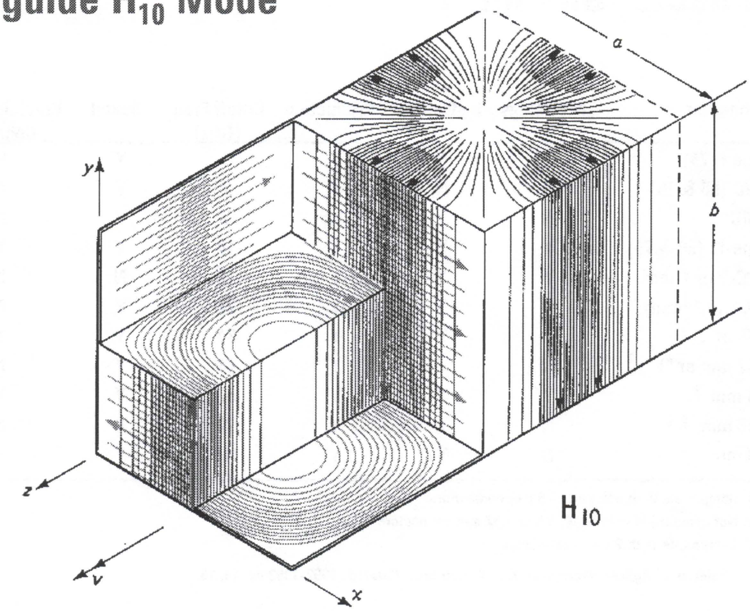
Agilent Technologies

Page 33

Frequency Coverage

- f max = approx. 120/D mm
- 7 mm = approx. 18 GHz
- 3.5 mm = 32 GHz
- Ratio D/d constant
- Depends strongly on dielectric support and mating pin geometry

Waveguide H_{10} Mode



Connector Summary

Connector	Metrology	Instrument	Production	Cutoff Freq (GHz)	Sexed	Precision Slotted Connector
Type F(75)	N	N	Y	1	Y	N
BNC (50 & 75)	N	N	Y	2	Y	N
SMC	N	Y	N	7	Y	N
Type-N (50 & 75)	Y	Y	Y	18	Y	Y
APC-7 or 7 mm	Y	Y	Y	18	N	N
SMA (4.14mm)	N	N	Y	22	Y	N
3.55 mm	Y	Y	Y	34	Y	Y
2.92 mm or "K" [1]	N	Y	Y	44	Y	N
2.4 mm [2]	Y	Y	Y	52	Y	Y
1.85 mm [2,3]	N	Y	Y	70	Y	N
1.0 mm	N	Y	Y	110	Y	N

1. Compatible with SMA and 3.5 mm connectors.
2. Not compatible with SMA, 3.5 or 2.92 mm connectors
3. Compatible with 2.4 mm connector

Reference: *Agilent Microwave Test Accessories Catalog* , 1992-1993 pp. 14, 15.

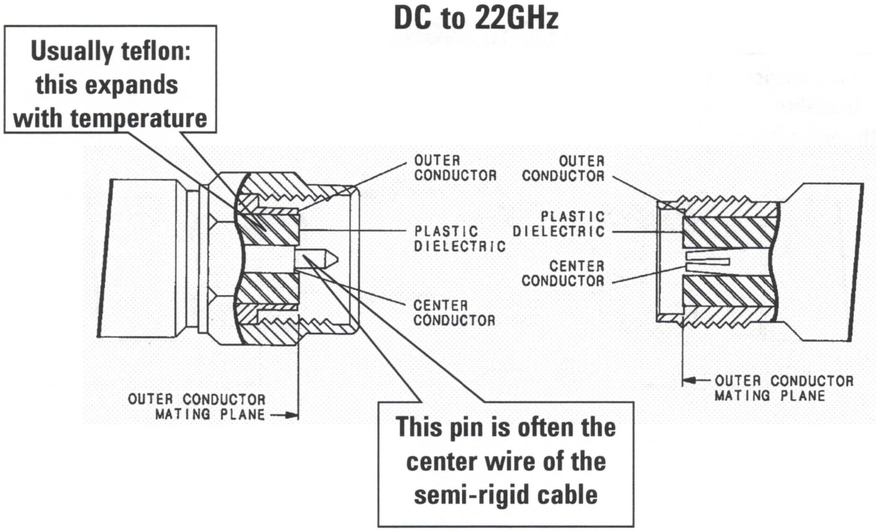
03 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 36

The SMA Connector

DC to 22GHz



03 – Transmission Lines
H7215A#101 v2.0

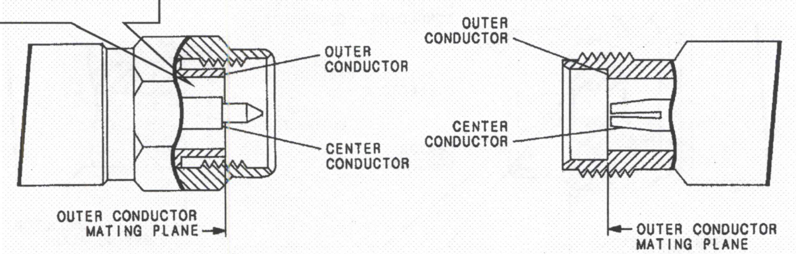
Agilent Technologies

Page 37

The Precision 3.5 mm Connector

DC to 34GHz

Air dielectric: is stable with temperature



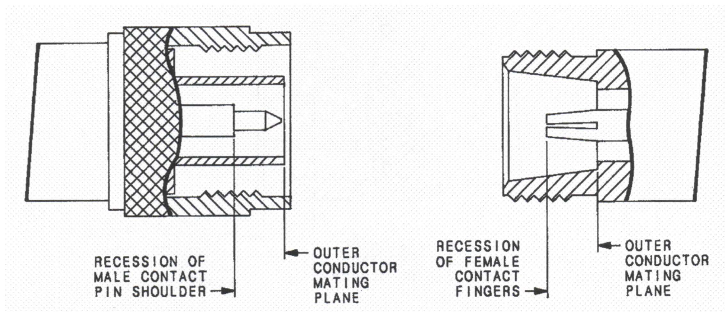
03 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 38

The Type N Connector

DC to 18GHz

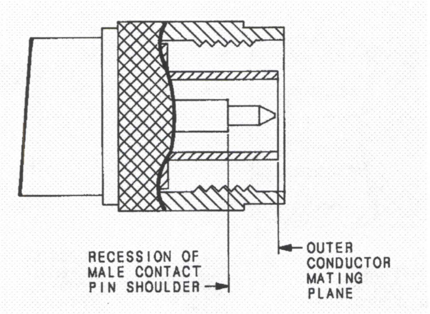


03 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 39

Pin Recession



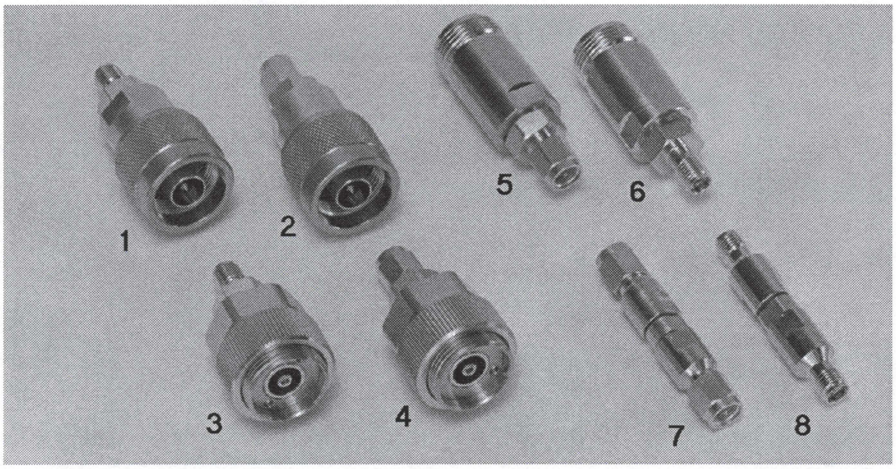
03 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 40

Some Precision Adapters

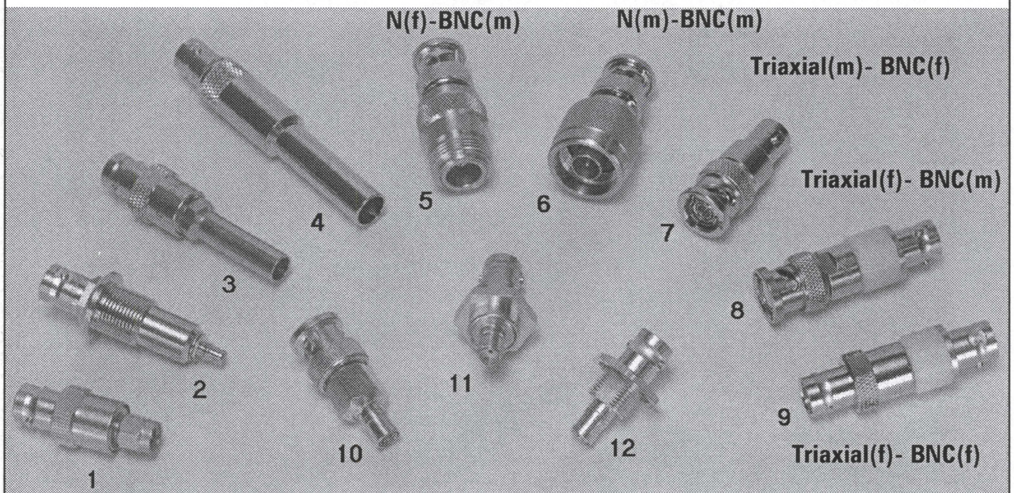
N(m)-3.5(f) N(m)-3.5(m) N(f)-3.5(m) N(f)-3.5(f)



7mm-3.5(f) 7mm-3.5(m) 3.5(m)-3.5(m) 3.5(f)-3.5(f)

Matched phase adapters

Some Precision Adapters

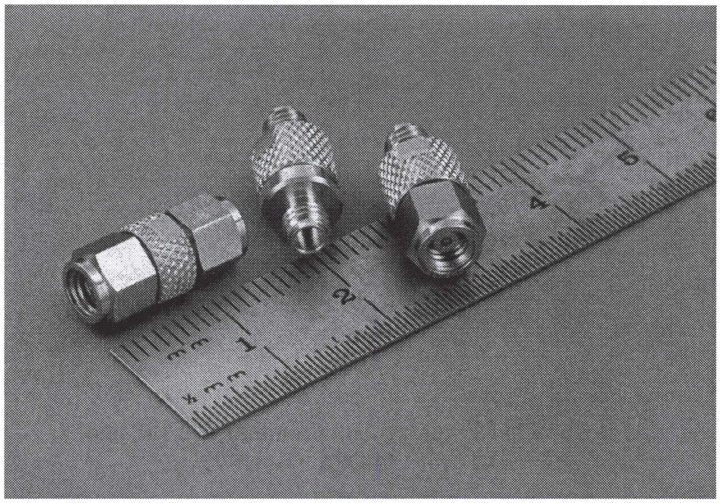


03 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 42

1.0 mm Coaxial Adapters



03 – Transmission Lines
H7215A#101 v2.0

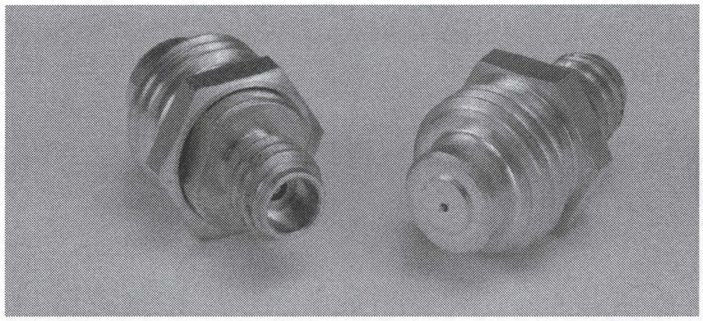
Agilent Technologies

Page 43

The 1.0 mm series adapters are designed to be used for the measurement of components with 50 ohm 1.0 mm connectors with a frequency range from dc to 110 GHz.

The 1.0 mm connector utilizes an air dielectric interface for the highest accuracy and repeatability. The coupling diameter and thread size maximize strength, increase durability, and provide highly repeatable connections. The connectors are designed so that the outer conductors engage before the center conductors.

1.0 mm Launch Adapter

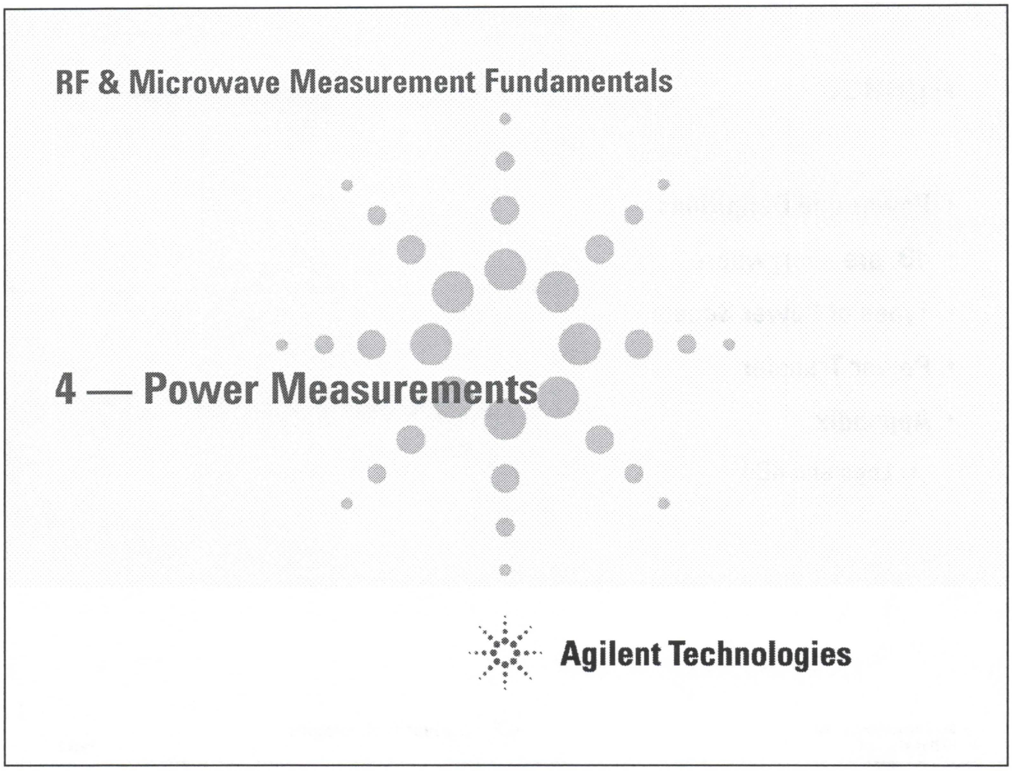


For Coax to Microstrip, pin diameter = 0.162 mm
11923A

03 – Transmission Lines
H7215A#101 v2.0

Agilent Technologies

Page 44



THE IMPORTANCE OF MICROWAVE POWER MEASUREMENTS

Introduction

Power measurements are related to key product specifications at every level in development, manufacturing, and quality control. Engineers involved in systems, such as radar or communication links, make power measurements to verify predicted performance. Components manufacturers are concerned with accurate characterization of their products with regard to heat and burnout. Quality control inspectors are responsible for determining which products meet published specifications. Power measurements must be made to verify those claims.

Overview

Power is the most basic measurement made at microwave frequencies. Average power is what is usually implied by the term "power".As transmission line theory has shown, our idea of incident and reflected waves may be extended to power flow, reflection and transfer. The transfer of power depends on the impedance of the source, load, and of the line connecting them.

The measurement of power is straightforward but subtle.

Agenda

- Power and Definitions
- dB, are everywhere
- Types of Power Sensor
- Power Transfer
- Appendix
  - Logs and dBs.

04 – Power Measurements
H7215A#101 v2.0

Agilent Technologies

Page 2

Why Voltage or Current is not Measured

- Transmission line effects.
- Probes can radiate.
- Probes can modify the impedance of the device.
- Waveguide and Coax do not allow easy probing.

04 – Power Measurements
H7215A#101 v2.0

Agilent Technologies

Page 3

At lower frequencies, from dc to several tens of megahertz, power can be measured simply by determining an impedance and measuring a voltage or a current. However, as frequencies go higher into the RF and microwave range, the measured values of voltage and current could change. As wavelengths become shorter, probes that measure voltage begin to radiate and impedances become extremely difficult to measure. In addition, transmission line effects may cause the voltage to vary with the position of the measurement on the line (remember, the voltage V_L anywhere on the line is Vi + Vr). Finally, the practical implementation of transmission lines, such as wave guide or coaxial cable, makes it extremely difficult to probe these voltages.

Why Power is Measured

- **Power is constant along a (lossless) line.**
- **Power is relatively simple to measure.**

04 – Power Measurements
H7215A#101 v2.0

Agilent Technologies

Page 4

The power flowing across any threshold on a lossless transmission line is not dependent on the position of that threshold. Power is relatively easy to measure.

Units of Power

- **Fundamental Units are:**
  - **Watt or**
  - **Joules/sec or**
  - **Horsepower, etc.**
- **1 Watt = 1 joule/sec**
- **746 watts = 1 hp**
- **Power is the rate of doing work**

04 – Power Measurements
H7215A#101 v2.0

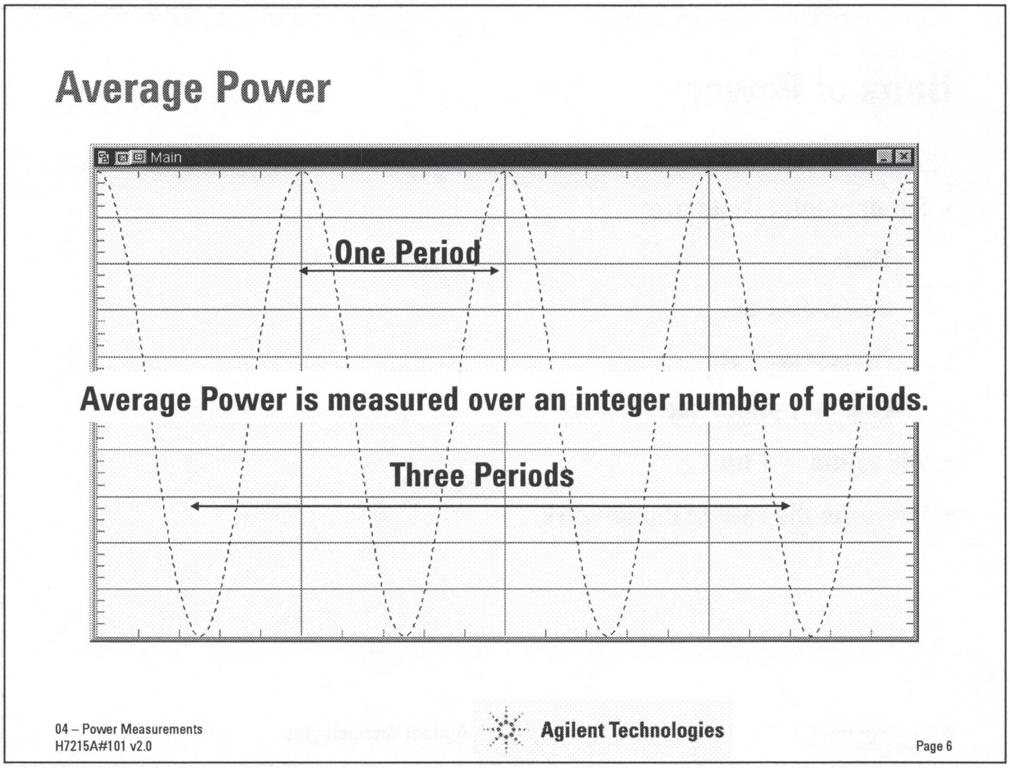
Agilent Technologies

Page 5

Types of Power

What is power?

Before discussing the specific techniques and devices used to make power measurements, we should first define the quantity that we are about to measure. Power can be defined simply as the rate of energy transfer. Sometimes, however, we are interested in specifying a power measurement for a defined interval of a waveform.

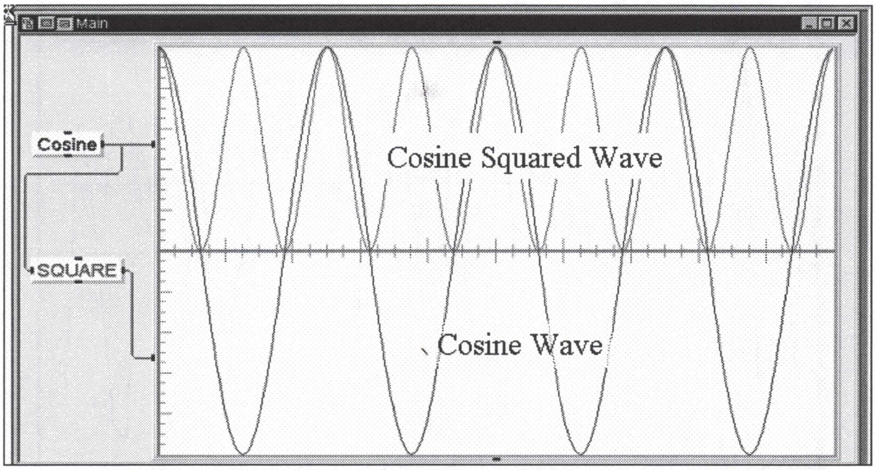


Average power

Average power is the quantity that is most frequently assessed during a power measurement. It is the energy transfer rate averaged over several cycles of the lowest frequency present in a waveform, (or at least an integer number of half cycles of a sine wave of constant amplitude). This type of measurement is often made by converting the energy to heat, and then determining the power by noting the temperature change in a sensing device.

The average value of the voltage wave shown is zero. We know that such a signal applied to a load will do work, so some other way to describe work done or power must be found.

Power ∝ V^2 (Square Law)



04 – Power Measurements
H7215A#101 v2.0

Agilent Technologies

Page 7

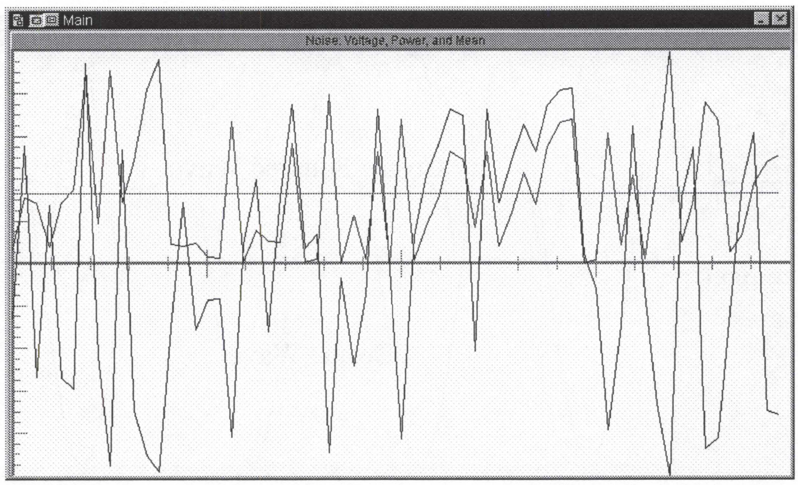
The relationship between power and voltage is shown here. With a resistive load the power is always positive and is proportional to V^2. If the output from a transducer is proportional to power it is a "square law" device. This plot shows instantaneous power. The average power for this condition is the mean of the instantaneous power calculated over at least one, or many integer number of cycles. This is easy to see for a sine wave. If the peak is ± 1 volt with a load resistance of 1 ohm, the instantaneous power will vary from 0 watt to 1 watt, the mean (average) can be seen to be 1/2 watt.

What is RMS?

A dc voltage of $\sqrt{(1/2)}$V = 0.7071V would produce the same average power in the load. To get this result we squared the voltage, found the mean, and then took a square root of the mean, this is the root-mean-square (rms) voltage. Thus for a sine wave the rms value is $\sqrt{(1/2)} \times$ peak voltage. Other waveforms have different rms values. By a similar argument an rms current can be found. Sometimes people refer to RMS power. This is bad science! What they probably mean is average power,which for a resistive load is:

$$= V_{rms} \times I_{rms}$$
$$= V^2_{rms} \div R$$
$$= I^2_{rms} \times R$$

Noise Waveform and Power



04 – Power Measurements
H7215A#101 v2.0

Agilent Technologies

Page 8

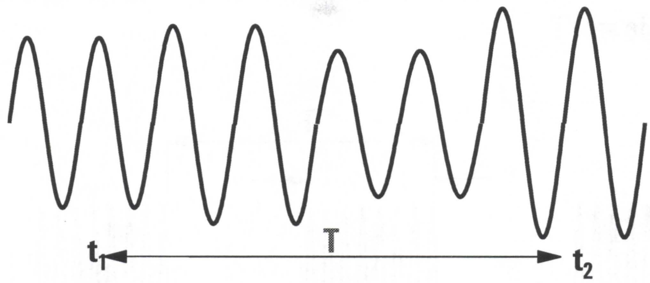
This represents a noise waveform with its associated instantaneous power. Over the time span represented, the rms value is 0.3327V (for this particular case). As noted earlier, the power is:

$$\frac{V_{rms}^{\ 2}}{R}$$

When measuring noise, the variation from reading to reading depends upon averaging time.

Average Power of a Changing Signal

- **The average power is related to a time period**

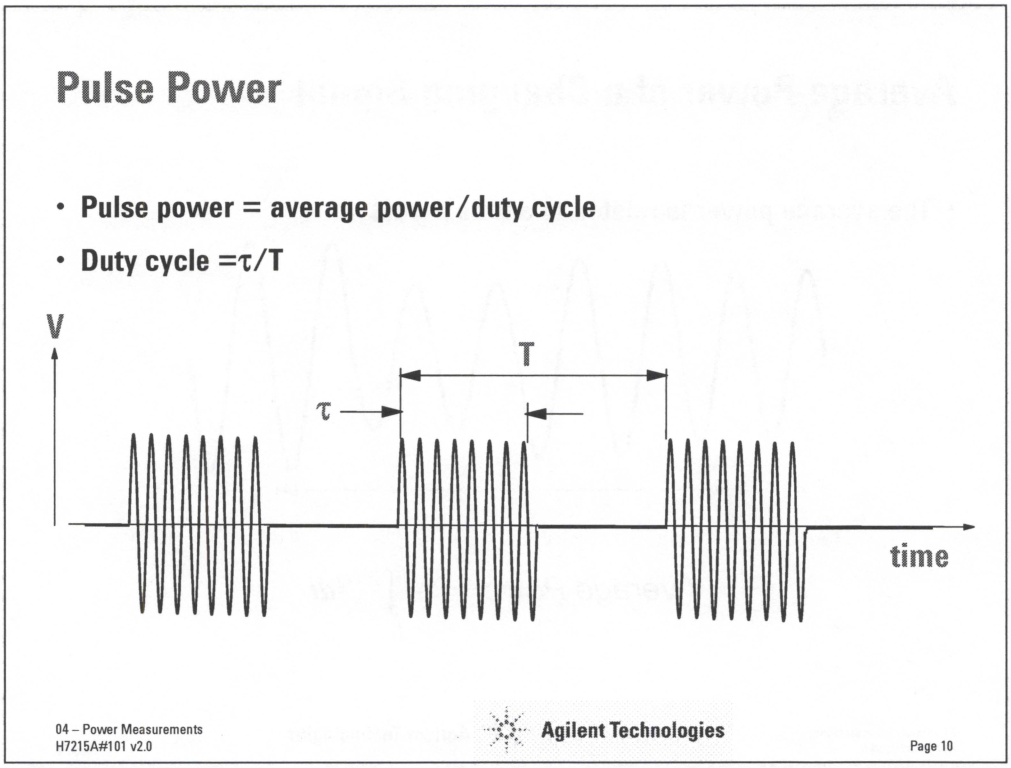


$$\text{Average Power} = \frac{1}{T}\int_{t_1}^{t_2}\frac{v^2(t)}{R}dt$$

04 – Power Measurements
H7215A#101 v2.0
Agilent Technologies
Page 9

The measurement of average power must depend on the time period over which the average is taken. So even an average power can vary if the wave amplitude is changing very slowly.

$$\text{Average Power} = \frac{1}{T}\int_{t_1}^{t_2}\frac{v^2(t)}{R}dt$$



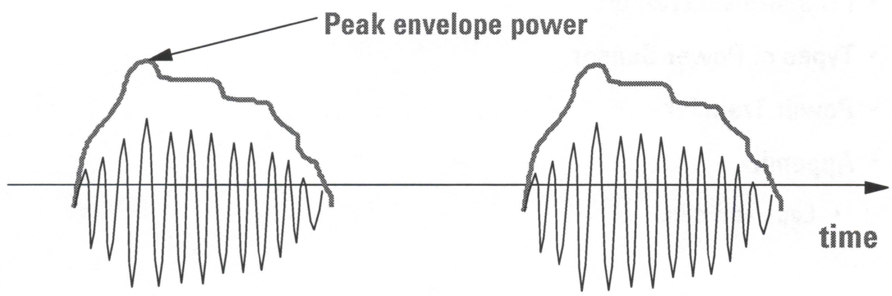
Pulse power

Pulse power measurements are often made in radar applications. The energy transfer rate is averaged over the pulse period. Any overshoot and ringing are also averaged. If the pulse is rectangular and the duty cycle is known, the pulse power is:

$$\text{Pulse power} = \frac{\text{Average Power}}{\text{Duty Cycle}}$$

Peak Envelope Power

- **Most pulse modulated RF is not so well behaved!**



Peak power

Peak envelope power is an average power (over a very short interval) used to describe the maximum envelope power of a signal. The particular wave shape is not important because the power is measured at the peak value of the modulation envelope. The averaging time must be less than $1/f_m$ where f_m is the maximum frequency component of the modulating waveform, but must be large enough to be many RF cycles long.

For perfectly rectangular pulses, the peak power and pulse power are equal to each other.

Agenda

- **Power and Definitions**
- **dB's, are everywhere**
- **Types of Power Sensor**
- **Power Transfer**
- **Appendix**
  - **Logs and dBs.**

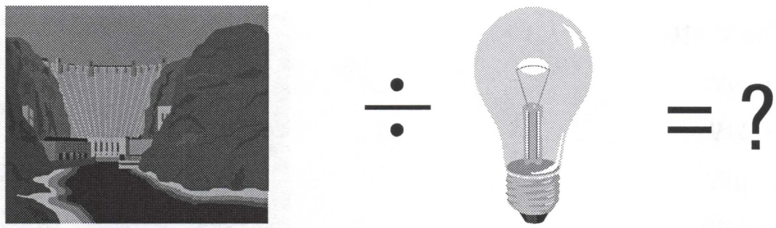
04 – Power Measurements
H7215A#101 v2.0

Agilent Technologies

Page 12

There follows a very brief outline about dBs and how they are used. In the appendix to the power measurement section is a dB tutorial with more detail.

Relative Power



- **We often need to express power in *relative* terms.**
  - **Signal to Noise Ratio**
  - **Sideband to Carrier Ratio**
- **Often expressed in logarithmic terms**
  - **decibel or dB.**

04 – Power Measurements
H7215A#101 v2.0

Agilent Technologies

Page 13

Units of Power

Although the International System (SI) unit of power is the watt, we will derive several other designations for use in this discussion. In many cases, such as when measuring gain or attenuation, the ratio of two powers, or relative power, is desired rather than absolute power. Relative power is typically expressed in decibels (dB).

$$\text{Power Ratio}_{\text{dB}} = 10\log_{10}\left(\frac{P_2}{P_1}\right)$$

Logarithms allow the expression of very large ratios and help in calculation of power level in different parts of systems.

A large power difference may exist in a transmitter/receiver system. The transmitter may have a output of 1kW , (103) W. The receiver may have a sensitivity of 1fW, (10--15W) Describing the difference in Watts or as a ratio is awkward. The use of logarithms make this task much easier. Since we usually deal with logarithmic terms let us be certain we understand and feel comfortable with logarithms before we proceed.

Absolute Power

- **The Watt**
  - **mW**
  - **kW**
  - **μW**
  - **MW**
- **dBm**
  - **dBμW**
- **Horse Power**

04 – Power Measurements
H7215A#101 v2.0

Agilent Technologies

Page 14

The absolute power as previously defined would have units of watts or a derivative of the watt. Here are a few multiples and submultiples of the watt.*

watt; W

megawatt; MW; 1,000,000 W; 10^6 W

milliwatt; mW; 0.001 W; 10^{-3} W

microwatt; μW; 0.000 001W; 10^{-6}W (Sometimes written uW) because some display devices do not support Greek characters.

picowatt; pW; 0.000 000 000 001W; 10^{-12}W

A dB based unit may be used for absolute power:

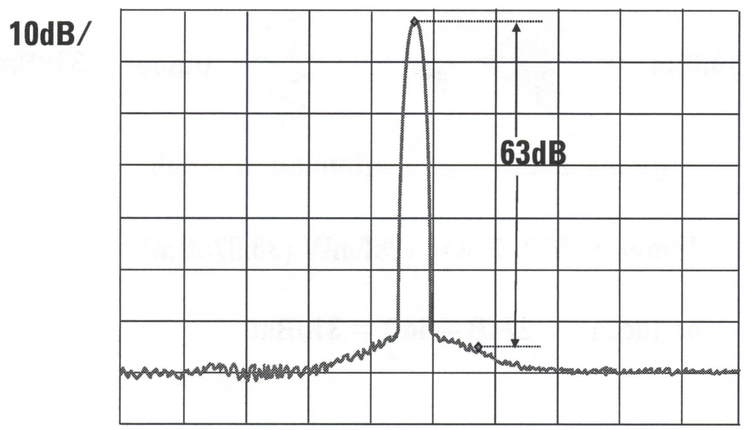
The dBm is based on one milliwatt and is defined: 10log(P/1mW)

$$\text{dB referenced to } P_{REF} = 10\log_{10}\left(\frac{P}{P_{REF}}\right)$$

The dBW and dBμW may be defined as 10log(P/1W) and 10log(P/1mW) respectively. P must be in the corresponding units.

**A more complete list of often used prefixes for all SI units is supplied in the dB appendix.*

dB's Express Large and Small Ratios



SNR = 63dB = 1:2,000,000 Power Ratio

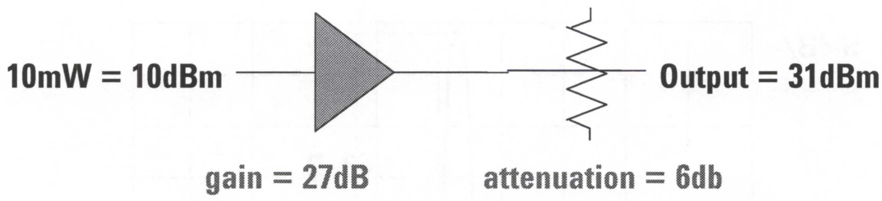
04 – Power Measurements
H7215A#101 v2.0

Agilent Technologies

Page 15

The main use use of the dB is the ability to easily express very large ratios. In the above example a displayed signal to noise ratio of 63dB would represent a linear ratio of 1:2000000.

dB's May Be Added and Subtracted



$10\text{mw} \times 500 \div 4 = 1250\text{mW}\ (30.97\text{dBm})$

or $10\text{dBm} + 27\text{dB} - 6\text{dB} = 31\text{dBm}$

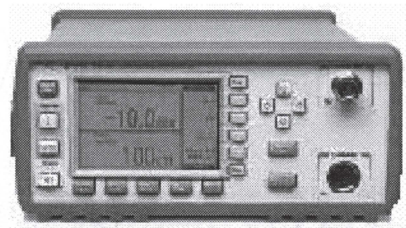
04 – Power Measurements
H7215A#101 v2.0

Agilent Technologies

Page 16

Here is a simple example of dB's in action. It is required to know how much power is available at the output of the circuit. In the above example an input of 10dBm undergoes amplification and attenuation to yield an output of 31dBm. This was easy to do, we have also indicated how the same result may be obtained by multiplying and dividing the linear quantities. The small discrepancy is due to the approximation made for the dB values. 27dB is linear 501.2 and 6dB is linear 3.98.

Agenda

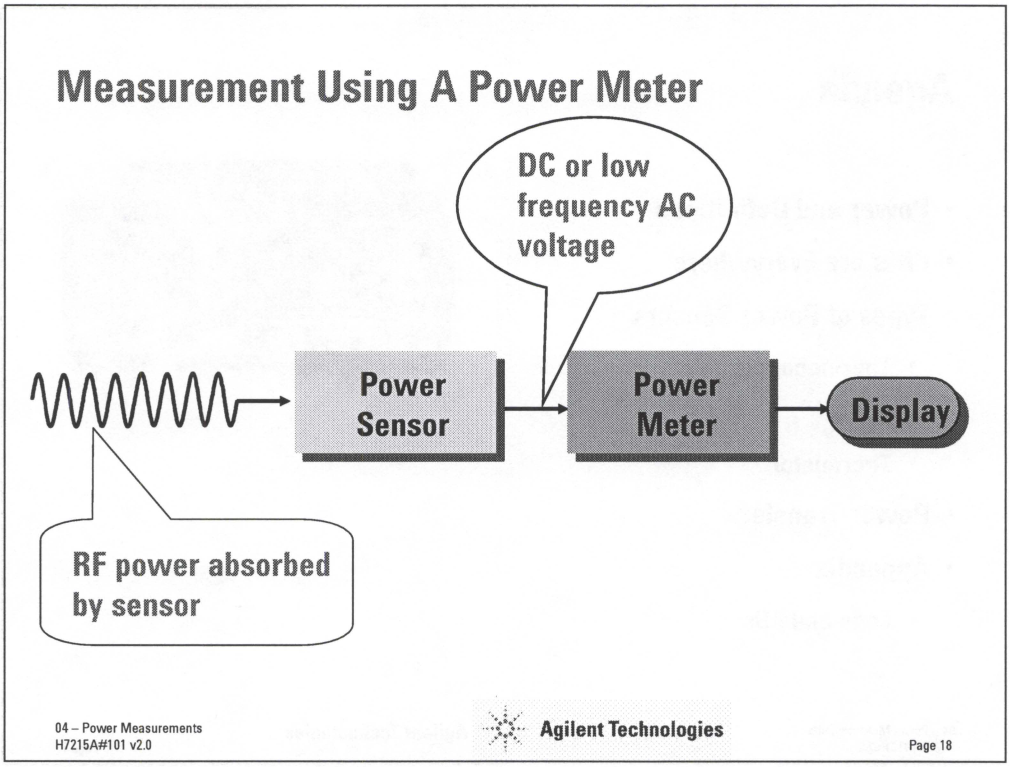


- **Power and Definitions**
- **dB's are Everywhere**
- Types of Power Sensors
    - **Thermocouple**
    - **Diode**
    - **Thermistor**
- **Power Transfer**
- **Appendix**
    - **Logs and dBs**

04 – Power Measurements
H7215A#101 v2.0

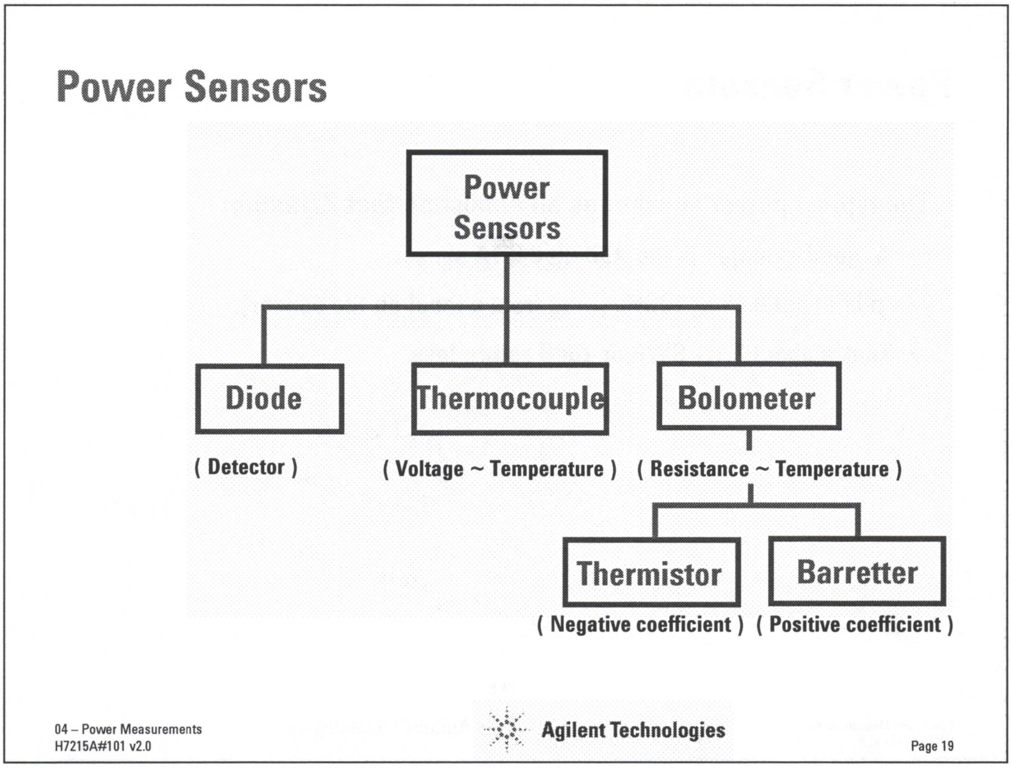
Agilent Technologies

Page 17



Let's now look at the types of hardware, both sensors and meters, that are used in average power measurements.

The basic idea behind a power sensor is to convert high frequency power to a DC or low frequency signal that the power meter can then measure and relate to a certain RF power level. The three main types of sensors are thermistors, thermocouples, and diode detectors. There are benefits and limitations associated with each type of sensor. We will briefly go into the theory of each type and then talk about the advantages and limitations associated with each sensor.



Sensors / Detectors

Families of sensors/detectors:

Power measurements usually are made by either of two basic techniques: diode detection, that makes use of diode rectification, or thermal detection, where the sensing device undergoes a proportional change in voltage, current, or resistance as a result of a change in temperature caused by the power.

Power Sensors

- The typical power sensor has an almost perfect Z_0 match.
  - A good example is the Agilent 8481A sensor.
  - ρ is about 0.05 at 10GHz (read from a label on the sensor).
  - That's about 1.1:1 SWR or 26dB return loss.

$\rho = \frac{\text{Reflected}}{\text{Incidental}}$

$\emptyset \rightarrow$ optimum $\Rightarrow$ No reflected

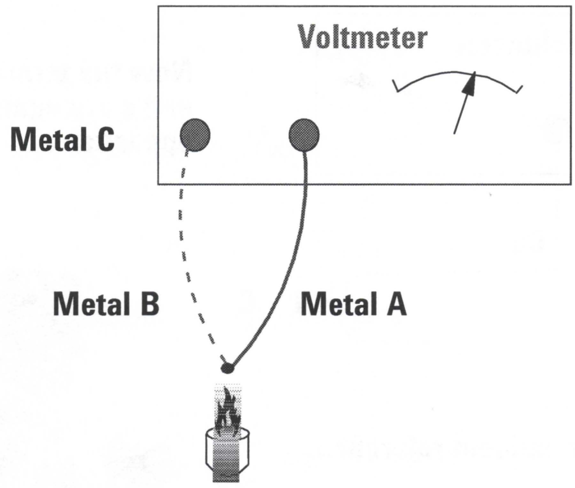
04 – Power Measurements
H7215A#101 v2.0

Agilent Technologies

Page 20

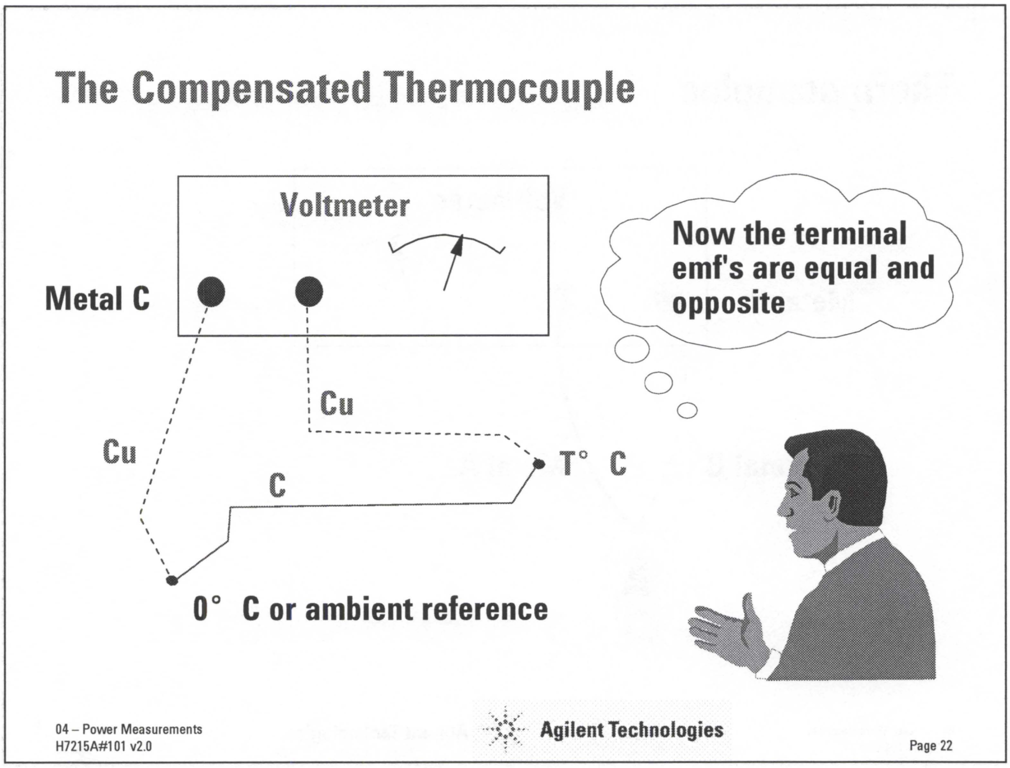
Power may be defined in several ways, in practical RF and microwave measurements power is measured on a Z_0 basis. Power sensors are designed to present an impedance of Z_0 to the device being measured, in order to give a value of the power that would be transferred to the transmission line basis of the system. The perfect Z_0 sensor is never actually achieved, but the reflections are very small.

Thermocouples



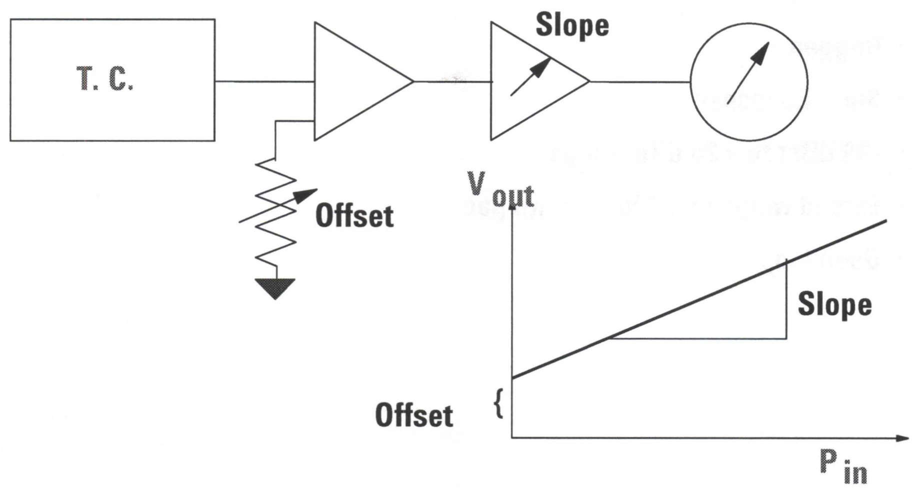
Thermocouples

Devices that change either voltage or resistance as a function of temperature lend themselves nicely to average power measurements. When power is incident on and is dissipated in these devices, the resulting rise in temperature changes the voltage or resistance in proportion to the dissipated power. The junction voltage of a thermocouple varies as a function of temperature.



Actually, thermocouple power sensors have two junctions, one maintained at room temperature (the metal case of the sensor), as a reference, and one exposed to the incident RF power to be measured This avoids the problem of unknown Seebeck voltages at the voltmeter terminals.

Calibrate and Zero for Open Loop Sensors



As RF power changes the temperature of the sensing junction, a net voltage difference, DV, occurs between the two junctions. With appropriate choice of junction materials, this voltage can be made a linear function of the dissipated RF power. All that is needed for accuracy are two known calibration points. These correct for the fixed dc offset at zero dissipated power, and the slope of the thermocouple response. Thermocouple power sensors can provide highly accurate power measurements from -30 to +20dBm. These have the lowest available reflection coefficient.

Thermocouple Characteristics

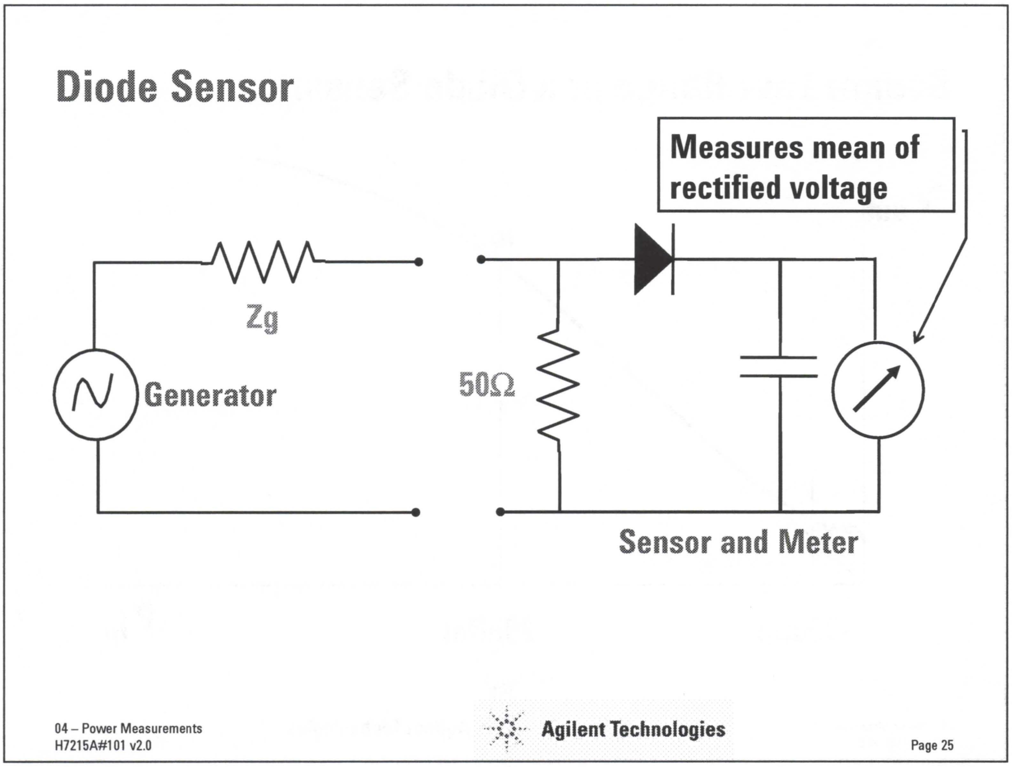
- Rugged
- Slow Response
- -30 dBm to +20 dBm range
- Extend range to +40dBm with pad
- Open loop

04 – Power Measurements
H7215A#101 v2.0

Agilent Technologies

Page 24

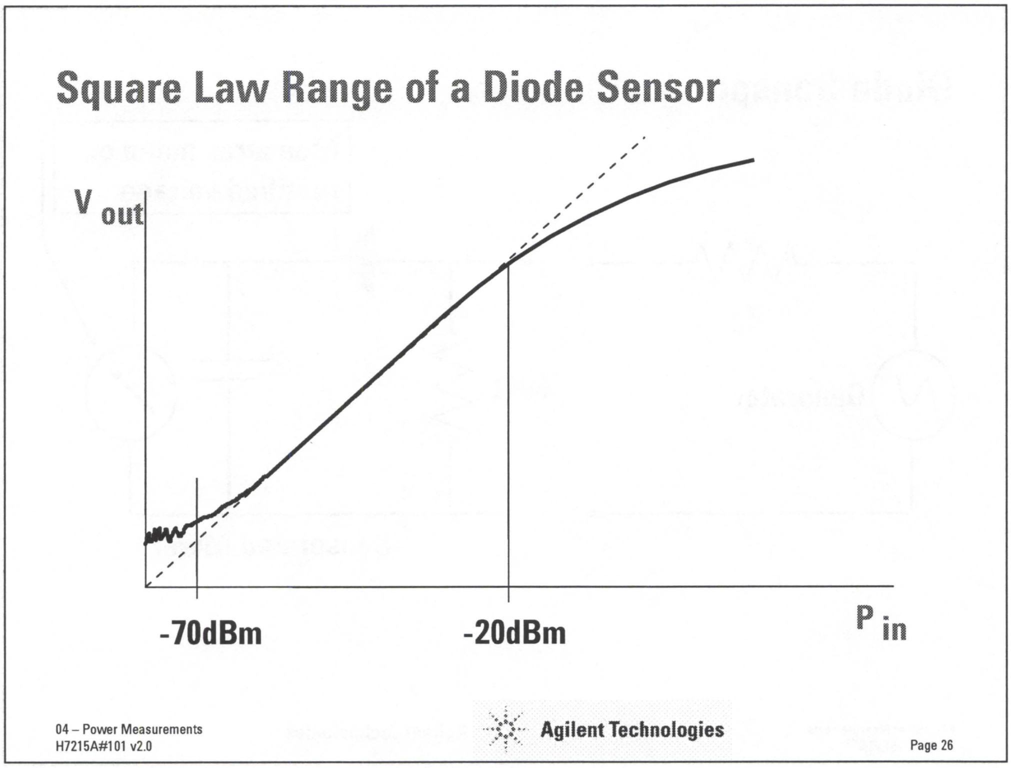
Measurement systems using thermocouples run open loop. To achieve accuracy, power meters using thermocouple power sensors and diode sensors contain both a sensor zeroing circuit and a calibrated source of RF power (1 mW at 50 MHz is common)



Diode Sensors / Detectors

An alternative to thermal detectors is the diode detector. The mechanism of power sensing is, however, quite different. Through rectification, a diode provides a dc current that is proportional to the square of the voltage across it, for a given voltage range. With the addition of several other components, a diode power sensor can be constructed such that the output voltage is proportional to the input power ($Vout = kVin^2$). Because of this relationship, the diode sensor is called a "square law detector." The range over which this relationship holds is called the square law range.

With careful design, a diode sensor can have sub-microsecond response time for measuring pulse power as well as average power. As with the thermal sensors, it is of paramount importance that the diode power sensor have an impedance close to Z_0. A non-zero reflection coefficient at the sensor effects power meter accuracy. The diode impedance is critical, so good control and repeatability of the diode manufacturing process is essential.

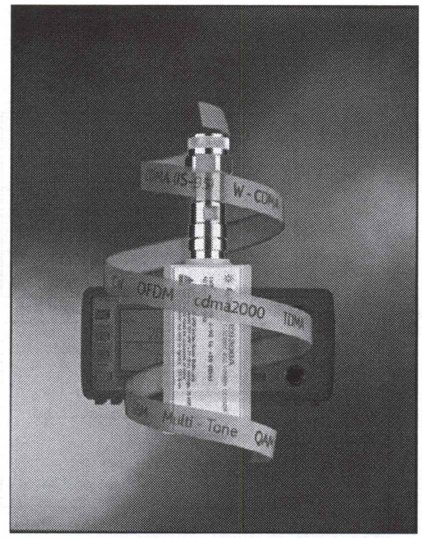


In a broadband diode power sensor, the lowest level measurable, known as the sensitivity, is about -70 dBm. This limit is simply the result of ambient thermal noise, which begins to corrupt the measurement at this point. The upper limit of the square-law range is -20 dBm. This provides a usable dynamic range of about 50 dB for multiple or complex signals as well as sine waves.

Some designs compensate for the range beyond square law and allow a total range from -70dBm to +20dBm, a 90dB dynamic range. Simple compensation will only allow accurate measurements of CW signals (sine waves) above the square-law range. Some more complex arrangements of multiple diodes allow the power measurement of modulated (wideband) signals over dynamic ranges up to 80dB.

Hallmarks of a Good Sensor

- **Sensor diodes always kept in square law region.**
- **Accurate measurement of signals with high peak-to-average ratios.**
- **Accurate measurement of signals with arbitrarily wide modulation bandwidth.**
- **Flat calibration factors give accurate measurement of multi-tone signals.**

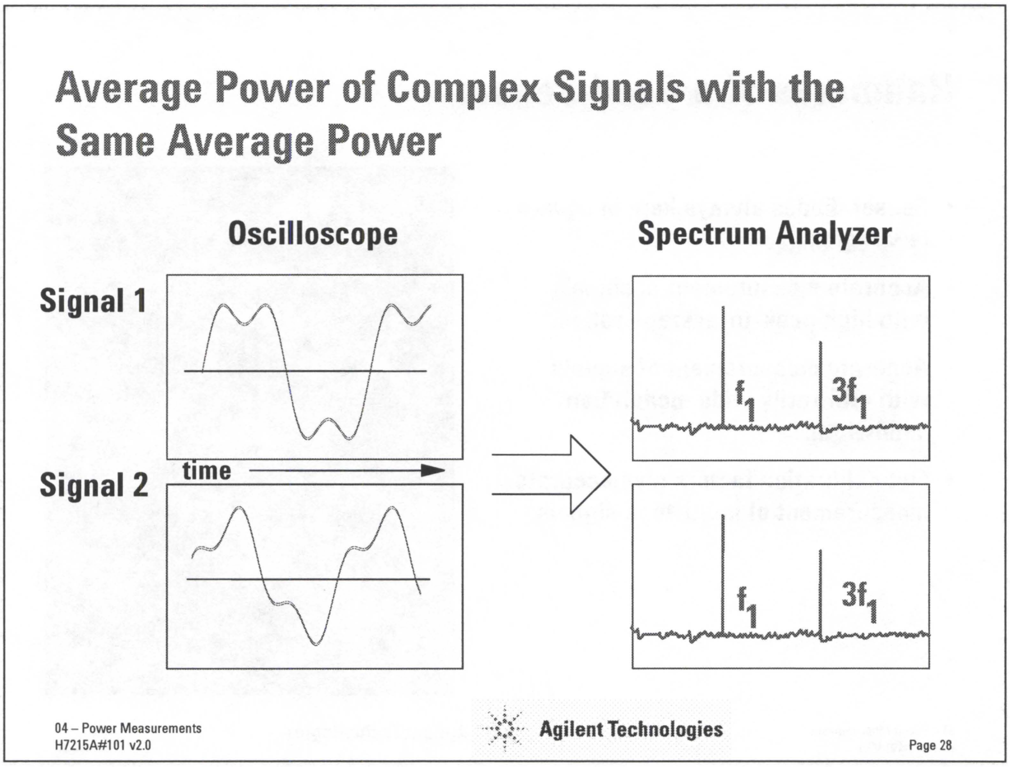


04 – Power Measurements
H7215A#101 v2.0
Agilent Technologies
Page 27

As mentioned the diode stack/attenuator/diode stack technique has the advantage of always using the diodes within their square law region. In this region, the output current (and voltage) of each diode is proportional to the input power. As a result, the E-Series E9300 power sensors respond accurately to measure average power on signals ranging from the most complex digital format to CW. This means that one E-series E9300 power sensor can be used to measure the average power of any signal.

The E9300 Power Sensor can handle power levels up to +33dBm peak with duration less than 10microseconds mean that these power sensors can be used for signals with high crest factors (or peak to average ratios). If the high peaks could not be handled this would effectively limit the dynamic range of the power sensor. The fact that it is a pure diode sensor also means that there is no modulation bandwidth limitations that you get with power sensors which use a sampling technique.

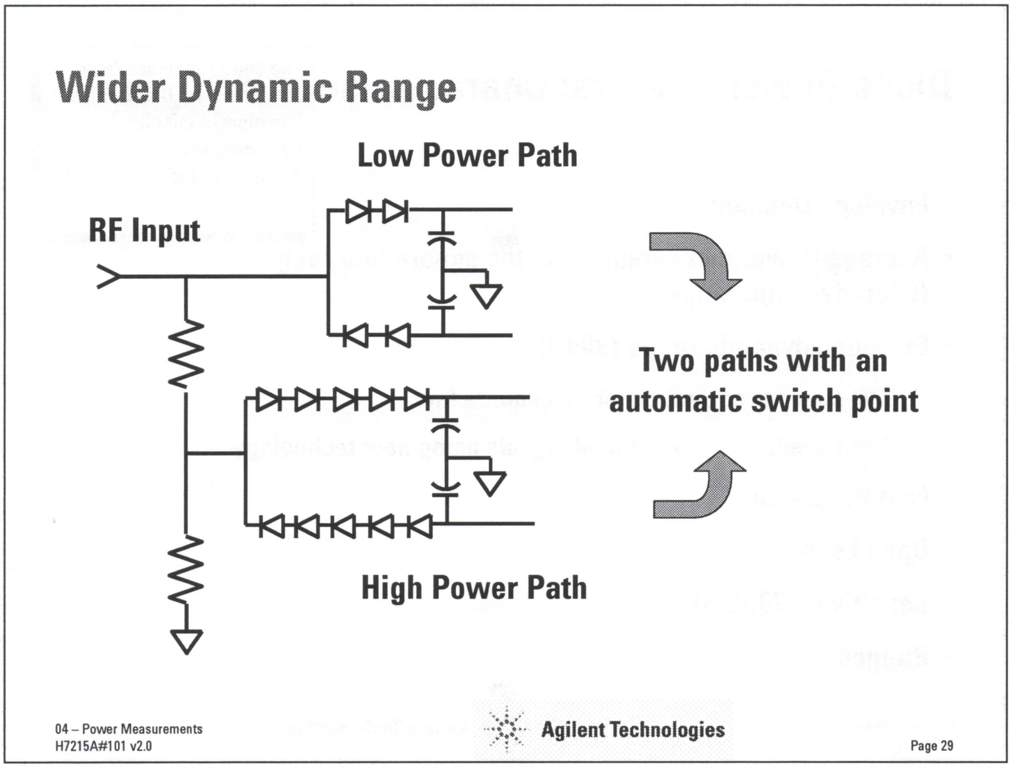
The E9300 Power Sensors have flat calibration factors which is an advantage if multi tone signals are being measured. If there are multi tone signals being input, only one calibration factor can be selected. If the calibration factors are flat over the frequency range of the measurements then the uncertainty in the result will be kept to a minimum.



Average power of complex signals

Complex signals have a phase relationship among their components. The slide shows two complex signals, both having two components. One component is at frequency f and the other component is a third harmonic at 3f. However, there is a different phase relationship between the components of the two signals which is only apparent when the two signals are viewed on an oscilloscope. In spite of this different phase relationship, the two signals have the same average power. The average power of a complex signal is the sum of the powers of the components as viewed in the frequency domain. This is the power that is measured with a thermocouple sensor and a diode sensor (in the square law range).

We say that these sensors are insensitive to phase. A diode sensor can also measure pulse power and peak power by teaming it with an appropriate power meter. As a square law detector, a diode sensor will provide a voltage proportional to any incident signal. The signal components need not be harmonically related. These signals may include noise. A simple diode sensor measuring such signals outside of the square law range would give incorrect and different results for the two signals shown here.



The corrections for the CW signals are fine but what can be done for other signals which may be modulated? The E9300 Power Sensors are diode power sensors which give 80dB Dynamic Range. To get this range a two path design is used with a separate path for the low power and high power paths. This innovative design is based on a diode stack /attenuator/diode stack topology.

Each diode stack forms a measurement path, the high power path is between -10 to +20dBm and the low power path is between -60 to -10dBm. Only one path is active at any time and switching between paths is fast, automatic and transparent to the user. This topology has the advantage of always maintaining the sensing diodes within their square law region and will therefore respond properly to complex modulation formats as long as the correct range is selected.

The design is further refined by incorporating diode stacks in place of single diodes to extend square law operation to higher power levels at the expense of sensitivity. In the E9300 the low power path uses a two diode stack pair and the high power path a five diode stack pair. FET switches were used off chip to enable the low path diodes to self bias to an off condition when not in use.

Diode Power Sensors: Characteristics

The 90dB range available for P sensors. Gigatronics has algorithmically corrected extended dynamic range.

- **Envelope Detector**
- **Average Power measurement in the square law range (50dB dynamic range)**
- **Extended dynamic range (90dB)**
  - **CW only for correction for non square law.**
  - **80dB available for wideband signals using new technology**
- **Fast Response**
- **Open Loop**
- **Sensitive (-70dBm)**
- **Rugged**

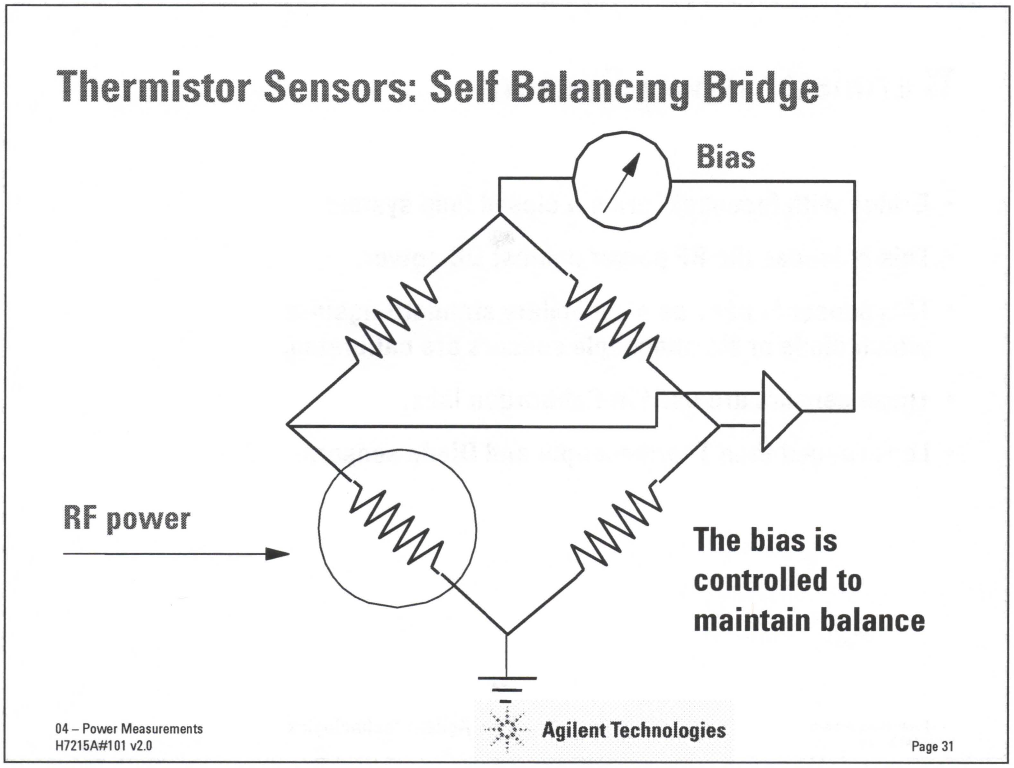
04 – Power Measurements
H7215A#101 v2.0

Agilent Technologies

Page 30

Diode sensors are sensitive and fast.

Thermocouple speed	(best with 437B, 438A)	20 readings/sec
Diode (E4412A, 4413A)	(best with EPM meters)	200 readings/sec



Thermistor sensors are used as a comparison to other power sensors. They are used by Calibration and Metrology Labs. These are the sensors that will be sent to a national lab (NIST, NPL etc.) for calibration. These calibrated sensors will then be used to standardize the *working* thermocouple and diode sensors.

Thermistor Power Sensors

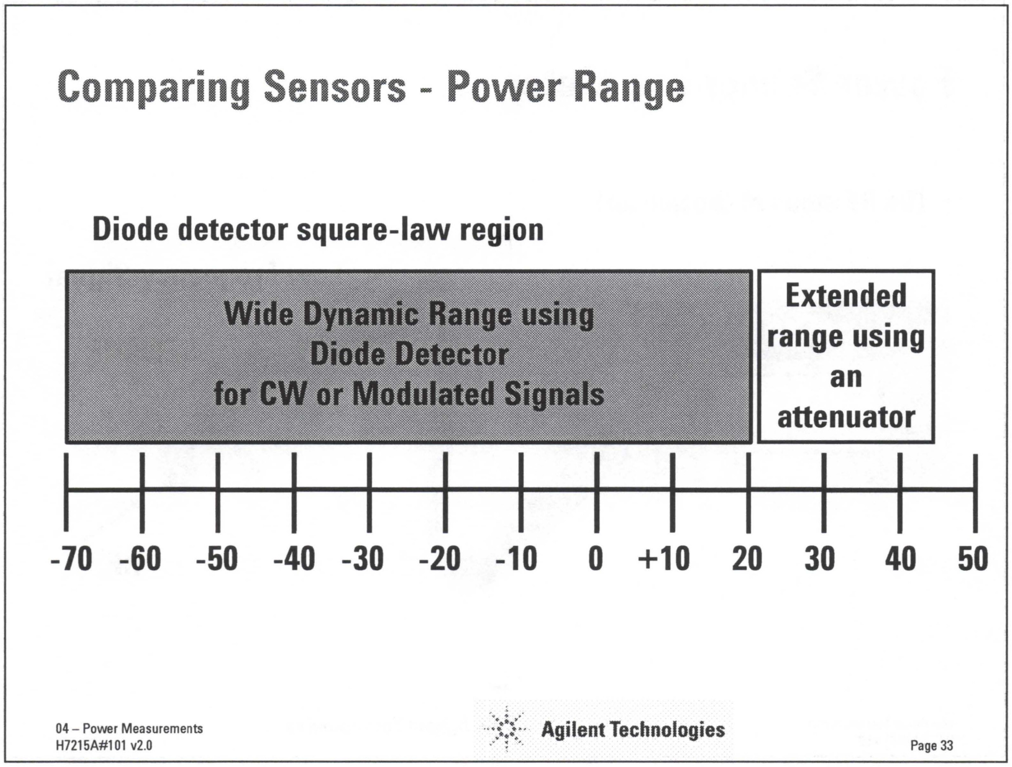
- Bridge with feedback forms a closed loop system.
- This balances the RF power against DC power.
- This sensor is used as a secondary standard against which diode or thermocouple sensors are calibrated.
- These sensors are used in Calibration labs.
- Less rugged than Thermocouple and Diode sensors.

04 – Power Measurements
H7215A#101 v2.0

Agilent Technologies

Page 32

Thermistor standards are secondary standards, and operate on the basis of substitution power. They are no longer found in general use because of their fragility compared to the thermocouple and diode sensors.



Thermistors offer high accuracy, but have a more limited operating range than a thermocouple or diode detector sensor. Thermistor mount specifications are for the range from -20 dBm to +10 dBm.

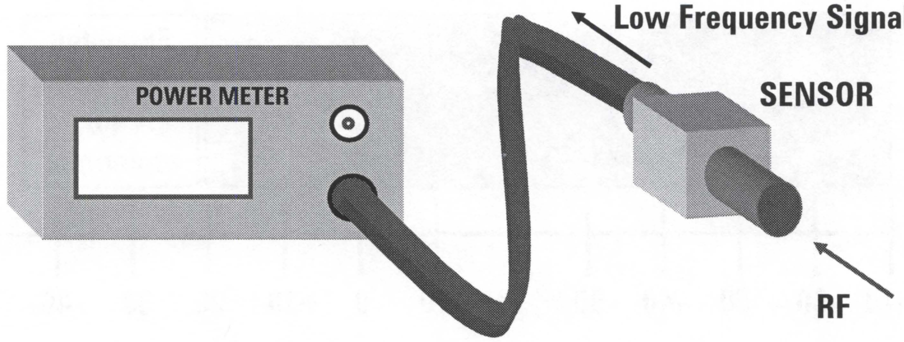
Thermocouples cover a very large range of powers. Their true square-law region is from -30 dBm to +20 dBm, and with an attenuator can operate up to +44 dBm. Three families of thermocouple sensors cover the complete -30 to +44 dBm range. The A-Series covers -30 to +20 dBm, the H-Series covers from -10 to +35 dBm, and the B-Series covers from 0 to +44 dBm.

Diode detectors (D-Series) have the best sensitivity, allowing them to work well below -20 dBm (stated range is -70 to -20 dBm), but above -20 dBm they begin to deviate substantially from the square-law detection region.

The Wide Dynamic Range Power Sensors are diode sensors and can provide up to 90dB dynamic range. They either work by correcting for the deviation (CW Power Sensors) or by using the two path technique to allow modulated measurements. Wide Dynamic range measurements can be made up to a maximum power of +44dBm.

Power Sensor and Meter

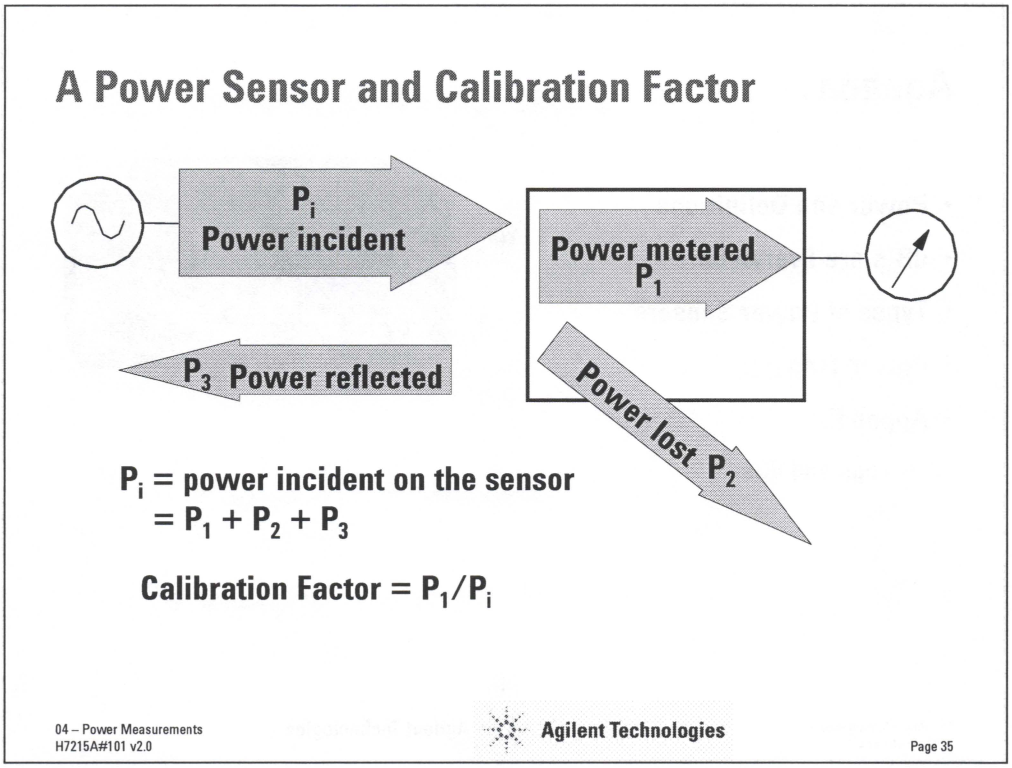
- **The RF stops at the sensor!**



04 – Power Measurements
H7215A#101 v2.0

Agilent Technologies

Page 34



Making a Power Measurement

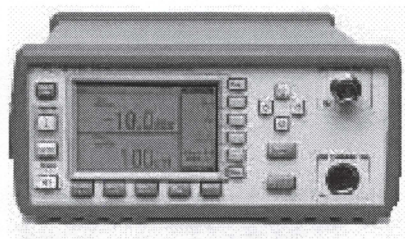
Practical power sensor

In measuring power, all incident power should be dissipated and metered by the measurement system. However, not all of the power will be dissipated in the sensor and properly metered. Some fraction actually is lost in other areas, perhaps dissipated in the walls and dielectric of the sensor. In addition, because the sensor itself is not a perfect Z0 load, it reflects some power which returns to the generator.

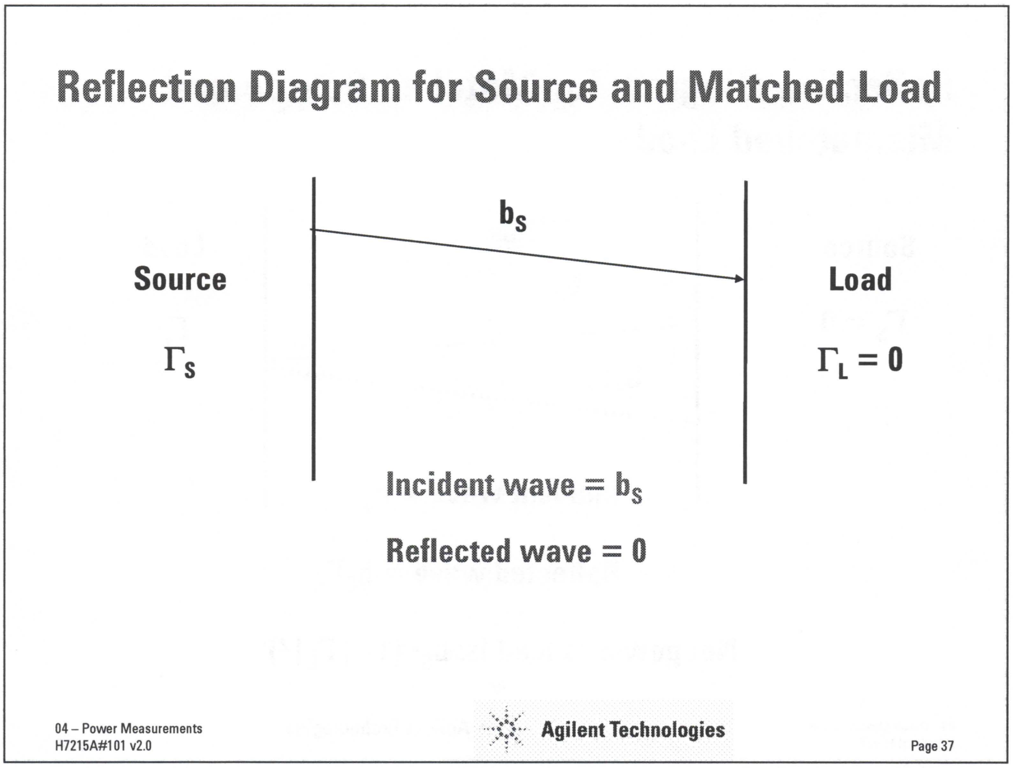
Calibration factor

Although power is lost because of the nature of the sensor, the loss is predictable and repeatable at any given frequency. It is a simple matter to offset and correct the final power meter reading by an appropriate amount. Calibration Factor is the percentage of the incident power that is actually converted for metering. It accounts for the mismatch loss caused by reflection from the sensor and the dissipated loss. Agilent power meters all provide external user adjustment of the Cal Factor based on a predetermined sensor response.

Agenda



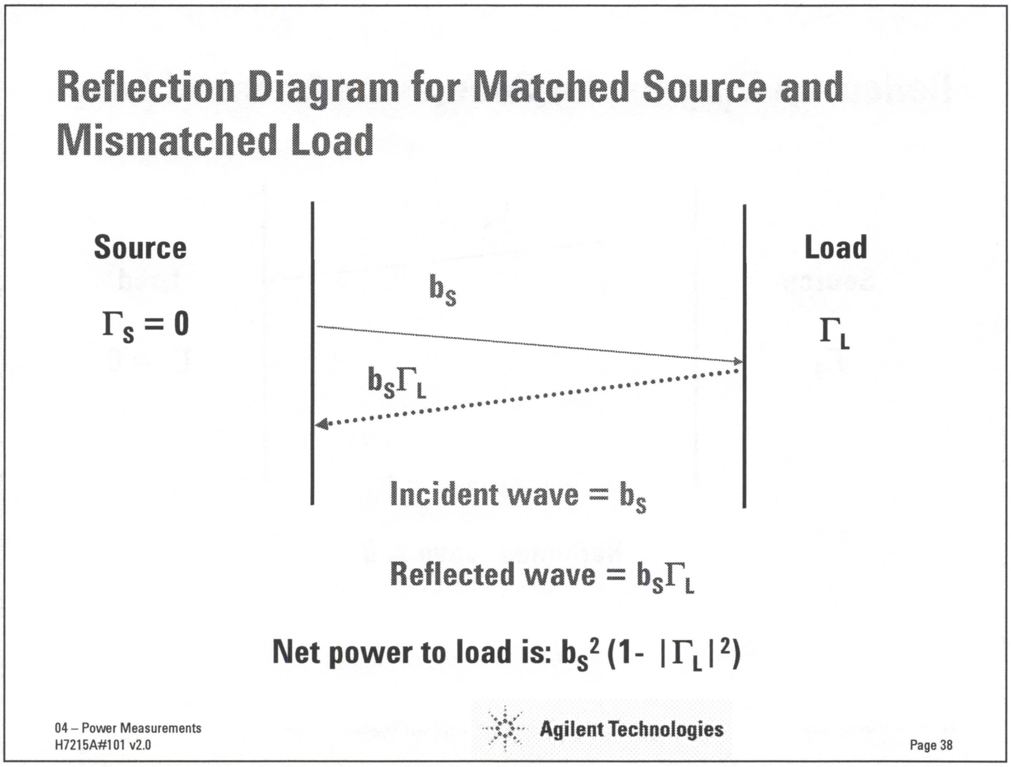
- Power and Definitions
- dB's are Everywhere
- Types of Power Sensors
- Power Transfer
- Appendix
  - Logs and dBs



To develop the expressions that describe the way in which power is transferred from a source through a transmission line to a load, we shall use a reflection diagram. The picture above is the ideal example with the incident wave being dissipated in the load. This corresponds to a power measurement on a **Zo basis.**

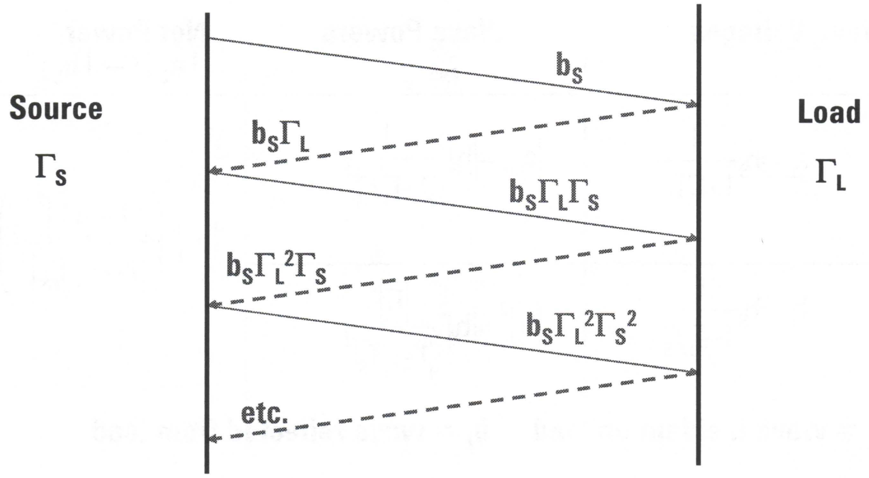
Notice that we use b_S as a reference, the source impedance will not be important until we analyze the effect of the re-reflected wave.

In the subsequent arguments the emerging waves are designated (a) and the incident waves as (b) so in this case the b_S is the wave emerging from the source and in later development becomes part of the wave incident on the load (a_L). The wave units, b and a are chosen so that $|b|^2$ and $|a|^2$ is power in watts.



If the load is *not* the same impedance as the characteristic impedance of the transmission line, then there will be a reflected wave, the magnitude and phase of this wave given by $b_s\Gamma_L$. In this case if the source is Zo so there will be no re-reflected wave.

Reflection Diagram for Mismatched Source and Load



04 – Power Measurements
H7215A#101 v2.0

Agilent Technologies

Page 39

Re-reflection from the source compounds the situation, but the calculation is quite simple. Adding all the forward waves and all the reflected waves in voltage gives a converging infinite series.

Σ(Incident waves) = $a_L = \ b_S + b_S\Gamma_L\Gamma_S + b_S(\Gamma_L\Gamma_S)^{2} + ... = b_S/(1 - \Gamma_L\Gamma_S)$

Σ(Reflected waves) = $b_L = \ b_S\Gamma_L + b_S\Gamma_L{}^2\Gamma_S + b_S\Gamma_L{}^3\Gamma_S{}^{2} + ... = b_S\Gamma_L/(1 - \Gamma_L\Gamma_S)$

Remember

$$\mathbf{1 + x + x^2 + x^3 + \ldots = 1/(1 - x)}$$

if $x < 1$

Incident and Reflected Voltage and Power

Wave Voltages	Wave Powers	Net Power $\lvert a_L \rvert^2 - \lvert b_L \rvert^2$
$a_L = b_S \frac{1}{1-\Gamma_L\Gamma_S}$	$\lvert a_L\rvert^2 = \lvert b_S\rvert^2 \frac{1}{\lvert 1-\Gamma_L\Gamma_S\rvert^2}$	$\lvert b_S\rvert^2 \left(\frac{1-\lvert\Gamma_L\rvert^2}{\lvert 1-\Gamma_L\Gamma_S\rvert^2} \right)$
$b_L = b_S \frac{\Gamma_L}{1-\Gamma_L\Gamma_S}$	$\lvert b_L\rvert^2 = \lvert b_S\rvert^2 \frac{\lvert\Gamma_L\rvert^2}{\lvert 1-\Gamma_L\Gamma_S\rvert^2}$	

a_L = wave incident on load **b_L = wave reflected from load**

Note: power towards load depends on the load!

04 – Power Measurements
H7215A#101 v2.0

Agilent Technologies

Page 40

The forward and reflected powers are the squares of the a and b wave quantities, so the net power delivered to the load may be expressed as

$$\lvert a_L \rvert^2 - \lvert b_L \rvert^2$$

Net power =

$$\lvert b_S\rvert^2 \left(\frac{1-\lvert\Gamma_L\rvert^2}{\lvert 1-\Gamma_L\Gamma_S\rvert^2} \right)$$

When the Gammas become rhos then to account for the maximum range that rho could have , the expression must have +/- rho

The Z_0 Mismatch Loss

Net power to actual load = $|B_s|^2 \left\{ \dfrac{1 - |\Gamma_L|^2}{|1 - \Gamma_L \Gamma_S|^2} \right\}$

Net power to perfect load ($\Gamma_L = 0$) = $|B_s|^2$

Z_0 mismatch loss = the *ratio* of these quantities expressed in dB

Z_0 mismatch loss in dB = 10 log $|1 - \Gamma_L \Gamma_S|^2$ - 10 log $(1 - |\Gamma_L|^2)$

$1 - \| \Gamma_L | ^2$

04 – Power Measurements
H7215A#101 v2.0

Agilent Technologies

Page 41

The mismatch loss is the ratio (Power that would be delivered to a Z_0 load) to (Net Power delivered to the actual load)

$\dfrac{|1-\Gamma_L\Gamma_S|^2}{1-|\Gamma_L|^2}$ which when expressed in dB is called "Z_0 mismatch loss"

$$=10\log|1-\Gamma_L\Gamma_S|^2 - 10\log\left(1-|\Gamma_L|^2\right)$$

when these reflection coefficients are known only as magnitudes

$$=10\log|1\pm\rho_L\rho_S|^2 - 10\log\left(1-\rho_L{}^2\right)$$

Γ_0

Γ_S

Uncertainty and Mismatch Loss with Scalar Reflection Coefficients

$$Z_0 \text{ mismatch loss in dB} = 10 \log |1 - \Gamma_L \Gamma_S|^2 - 10 \log (1 - |\Gamma_L|^2)$$

$$10 \log (1 \pm |\rho_L \rho_S|)^2 \qquad\qquad 10 \log (1 - \rho_L^2)$$

This part becomes the mismatch *uncertainty* if we don't know the phase associated with the gamma Γ. These become rhos ρ.

This part is the *mismatch loss.*

04 – Power Measurements
H7215A#101 v2.0

Agilent Technologies

Page 42

Z_0 Mismatch Loss = Mismatch Loss + Uncertainty (all in dB)

When the reflection quantities are known as scalars then the general mismatch loss expression becomes decomposed into an uncertainty and a loss term.

In a power sensor, because the loss term can be determined and only depends on the sensor match, it is combined with the other systematic quantities in the calibration factor.

This uncertainty, if associated with a power sensor measurement, is not the only uncertainty to be considered. See Agilent application note AN64-1A for a more complete description.

Note: On the Agilent Reflectometer Calculator, the mismatch uncertainty may be estimated from SWR1 and SWR2 on the "Mismatch Error Limits" side, use the other side to convert r to SWR.(SWR1 and SWR2 can be in any order order for SWRL and SWRS). The mismatch loss may be estimated from the "Reflectometer" side of the calculator.

Mismatch Uncertainty as a Fraction

Fractional change $\approx \pm\, 2\rho_L\rho_S$
$= \pm\, 2\rho_L\rho_S \times 100\%$

Mismatch uncertainty $= \pm\, 2\rho_L\rho_S \times 100\%$

04 – Power Measurements
H7215A#101 v2.0

Agilent Technologies

Page 43

From the uncertainty expression:

$$20\log(1\pm\rho_1\rho_2) = 10\log(1\pm\rho_1\rho_2)^2$$

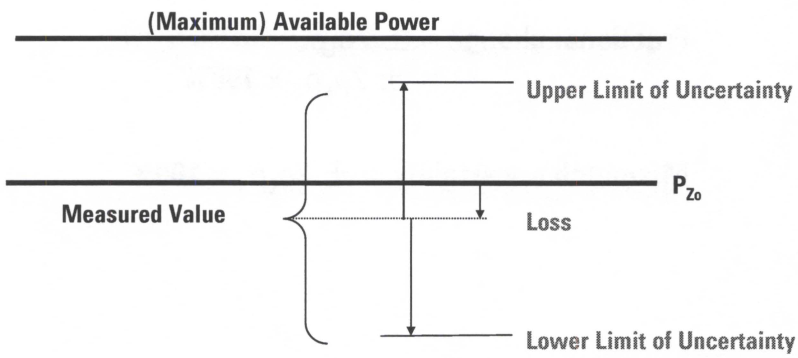
$$= 10\log(\textit{fractional range})$$

$$\textit{The fractional range} = (1\pm 2\rho_1\rho_2 + (\rho_1\rho_2)^2)$$

$$\textit{if } \rho_1\rho_2 << 1 \textit{ and neglecting } (\rho_1\rho_2)^2$$

$$\textit{the fractional change} \approx \pm 2\rho_1\rho_2$$

Power Measurement with Limits Due to Mismatches



The result of measurement could be anywhere between the uncertainty limits.

04 – Power Measurements
H7215A#101 v2.0

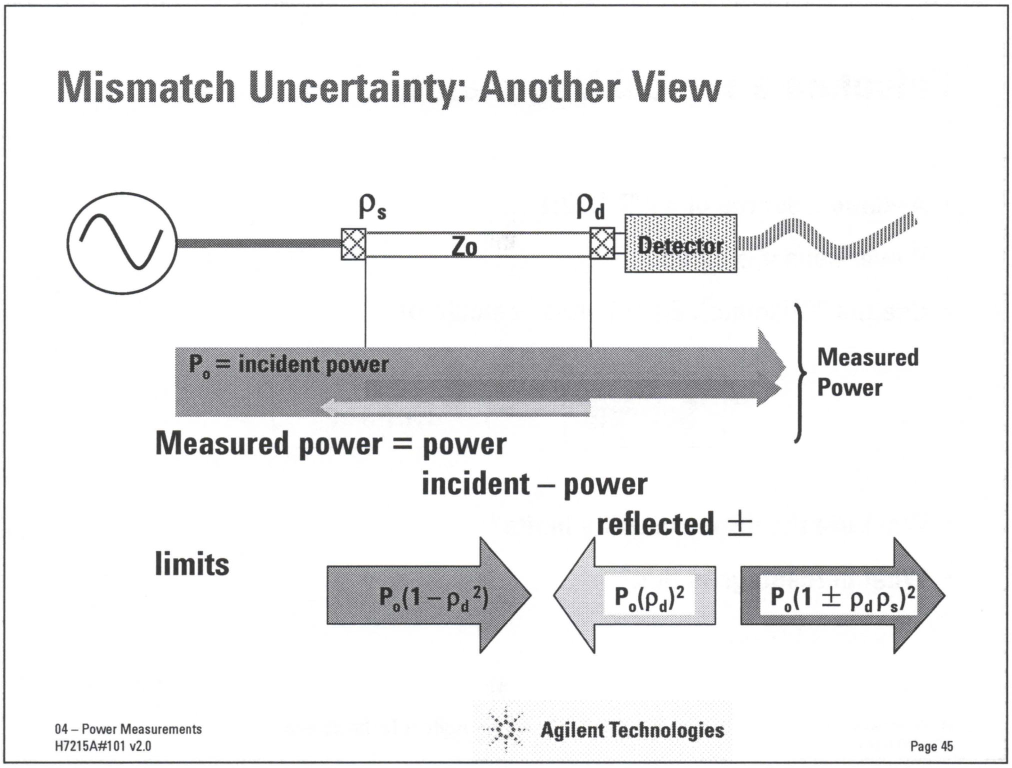
Agilent Technologies

Page 44

P_{Zo} is sometimes referred to as the "True Value" this is more properly called the measurand. This definition of power from a source is on a Z_0 basis, that is the net power that would be delivered to a Z_0 load (The power incident on the load from a Zo transmission line). The "source" could be any microwave device where the physical generator is separated from the load by a transmission line and/or other components. The actual characteristic impedance of the transmission line may or may not be Z_0, but Z_0 is the *basis* of the measurement.

Note that a measured value can be greater or less than P_{Zo}

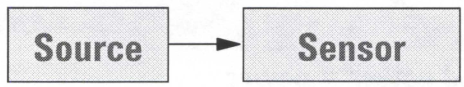
This is very evident when measuring insertion loss of a device with a nominal loss of zero, the power measured at the device output may be greater than the power measured at the device output.



The power sensor calibration coefficients include the effect of power reflected which is determined because ρd is measured when the sensor is manufactured and at subsequent re-calibration by a standards lab.

Calculate a Typical Uncertainty

- Assume a source of SWR 1.22:1
- Power sensor ρ of .05
- Use the "Mismatch Error Limits" calculator



- What are the dB uncertainty limits?
- What percentage is that?

Here is a problem to solve.

Power Measurement Uncertainty

- **From the calculator U = ± 0.04dB**
- **This is about ±1%**
- **This is considered excellent for a microwave measurement.**

04 – Power Measurements
H7215A#101 v2.0

Agilent Technologies

Page 47

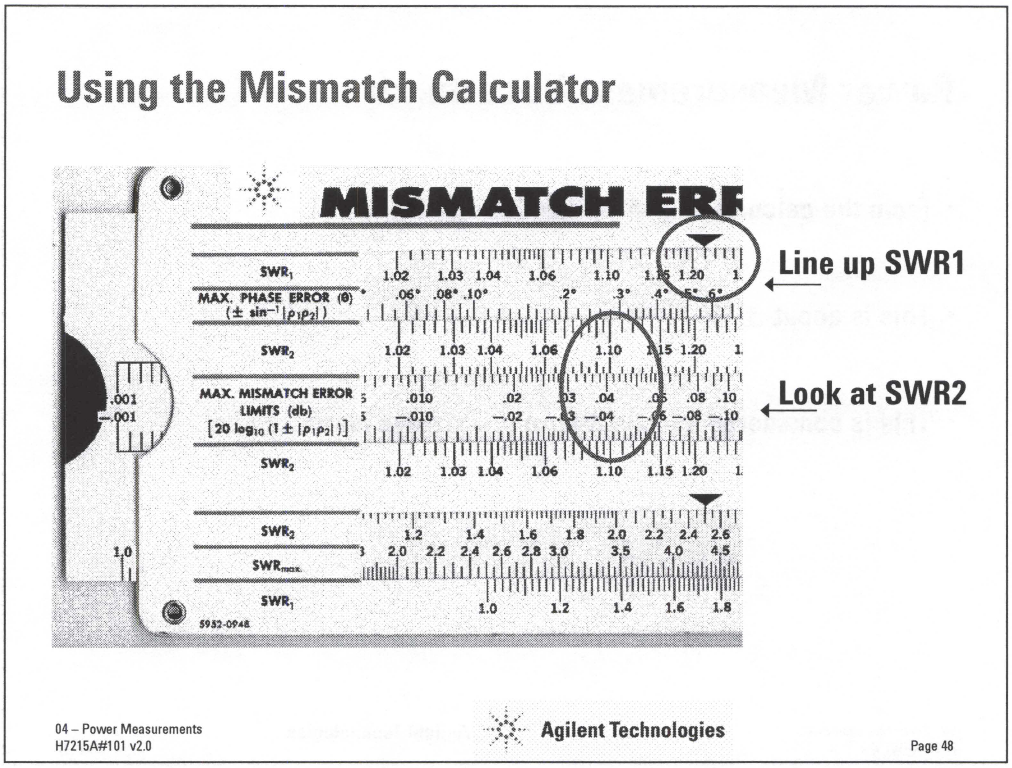
To work this problem using the formulae, first express the SWR as a reflection coefficient, using $\rho = (S - 1)/(S + 1)$.

$(1.22 - 1)/1.22 + 1) = 0.0991$

Using $U_{dB} = 20\log(1 \pm \rho_1\rho_2)$
Calculate the positive value
$20\log(1 + 0.099 \times 0.05) = + 0.043\text{dB}$

Calculate the negative value
$20\log(1 - 0.099 \times 0.05) = - 0.043\text{dB}$

For these small SWR values the positive and negative dB uncertainties are equal, to three decimal places.



To use the slide rule first flip to the reflectometer side to convert rho to SWR. Then use the mismatch error limits side to find the uncertainty as shown.

Other Sources of Power Meter Uncertainty

- **Uncertainties of:**
  - **Cal Factor**
  - **Ref. Oscillator**
  - **Ref. Oscillator mismatch**
  - **Instrumentation**
  - **$\pm$ 1 count**
  - **Zero Set**
  - **Zero Carryover**
  - **Noise**
  - **Drift**
  - **Linearity of sensor**

04 – Power Measurements
H7215A#101 v2.0

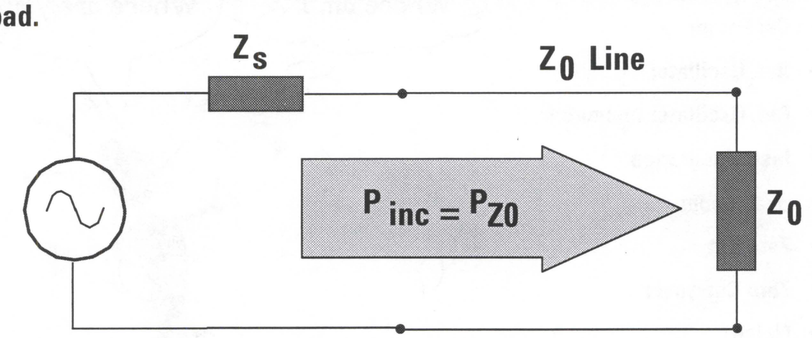
Agilent Technologies

Page 49

This lesson has concentrated on the RF and microwave aspects of power measurement and power transfer, where an analysis of power transfer reveals a large source of uncertainty, the mismatch uncertainty. When using a power meter there are other uncertainties to consider as listed above. It is beyond the scope of this class to evaluate the total uncertainty of a power measurement but an excellent analysis of the power meter is done in AN 64-1A.

Ideal Power Measurement

- We usually want to measure the power that would be absorbed by a Z_0 load.



Ideal Power Measurement

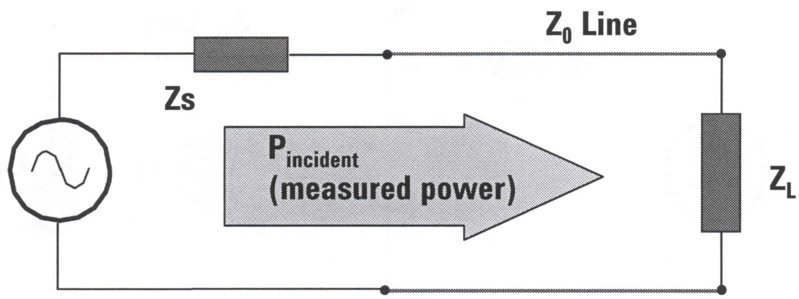
We now have some understanding of the problems in measuring RF and microwave power with imperfect generator and power sensor terminations. Let us summarize the various situations. P_{Z0} is measured or dissipated when the transmission line and termination are exactly Z_0 . P_{Z0} is the quantity we want.

Fundamental Goal

Every microwave signal source has some source impedance. If the impedance basis of the system is Z_0, typically 50 or 75 ohms, then we want to know how much power the source will deliver to a Z_0 load. The power delivered to the load is a function of the source impedance and the load impedance.

Practical Power Measurement

- **The required power lies between the limits of the mismatch uncertainty.**



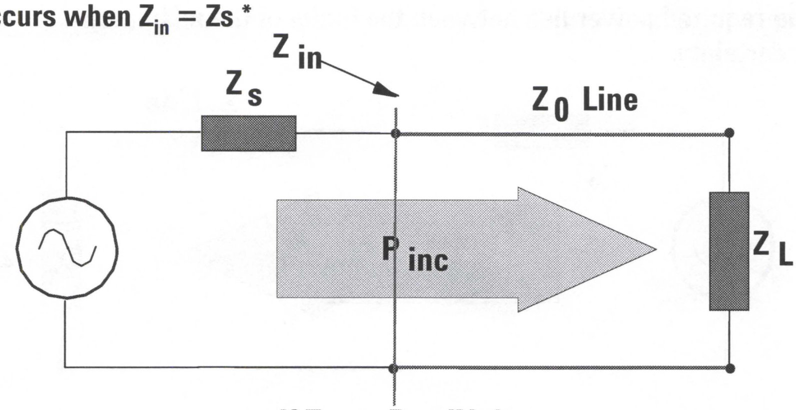
$$P_{inc}(1 - \rho_S \cdot \rho_L)^2 \leq P_{zo} \leq P_{inc}(1 + \rho_S \cdot \rho_L)^2$$

Practical Power Measurement

The incident signal includes both the Z_0 signal and the re-reflected signal.

Maximum or Available Power

- Occurs when $Z_{in} = Zs^*$



If $Z_S = R + jX$ then

$Z_S^* = R - jX$

04 – Power Measurements
H7215A#101 v2.0

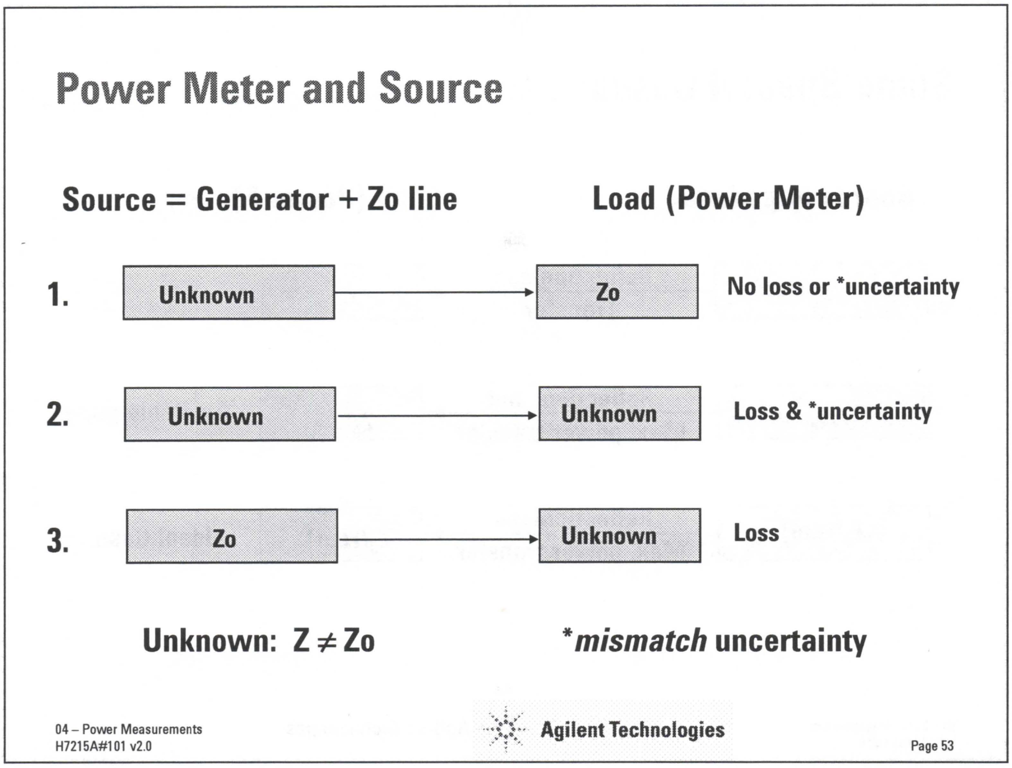
Agilent Technologies

Page 52

Maximum Power Transfer

We define Available Power or Conjugate Power when $Z_{IN} = Z_S^*$. It can be shown that when $Z_{IN} = Z_S^*$, then $\Gamma_{IN} = \Gamma_S^*$. This condition implies that the maximum power possible is being dissipated in a load.

In an ideal system Z_0 is real (lossless line) and $Z_S = Z_0 = Z_L$ then $P = P_{Zo}$

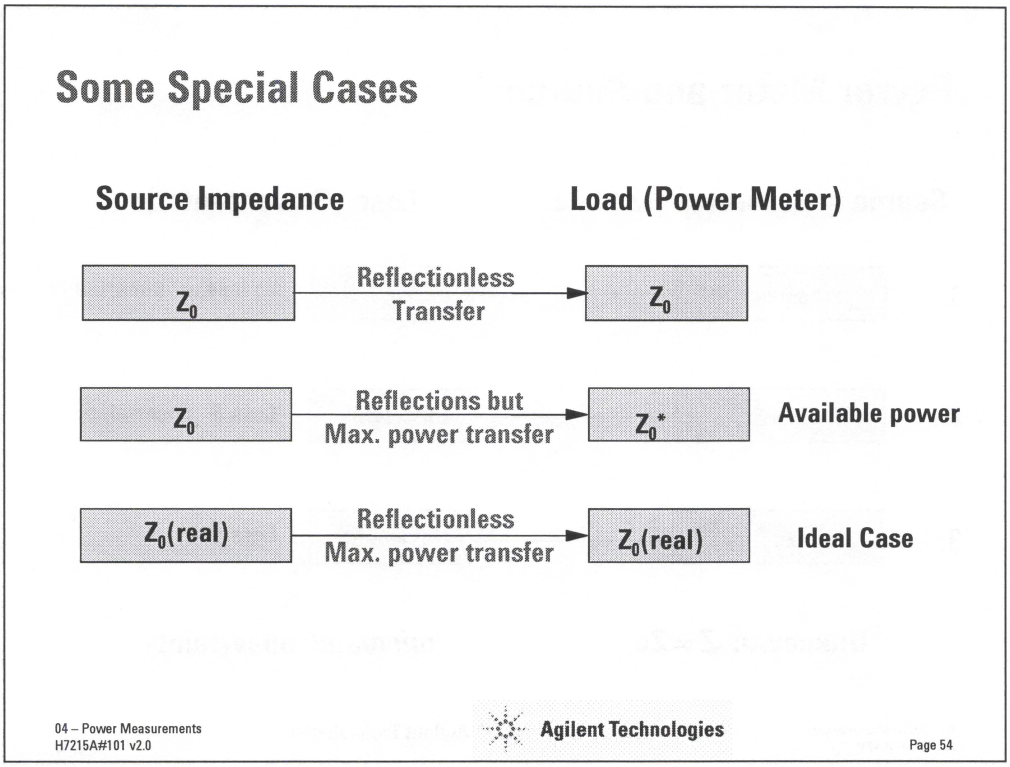


Remember that we want to know what power a source will deliver to a Z_0 load.

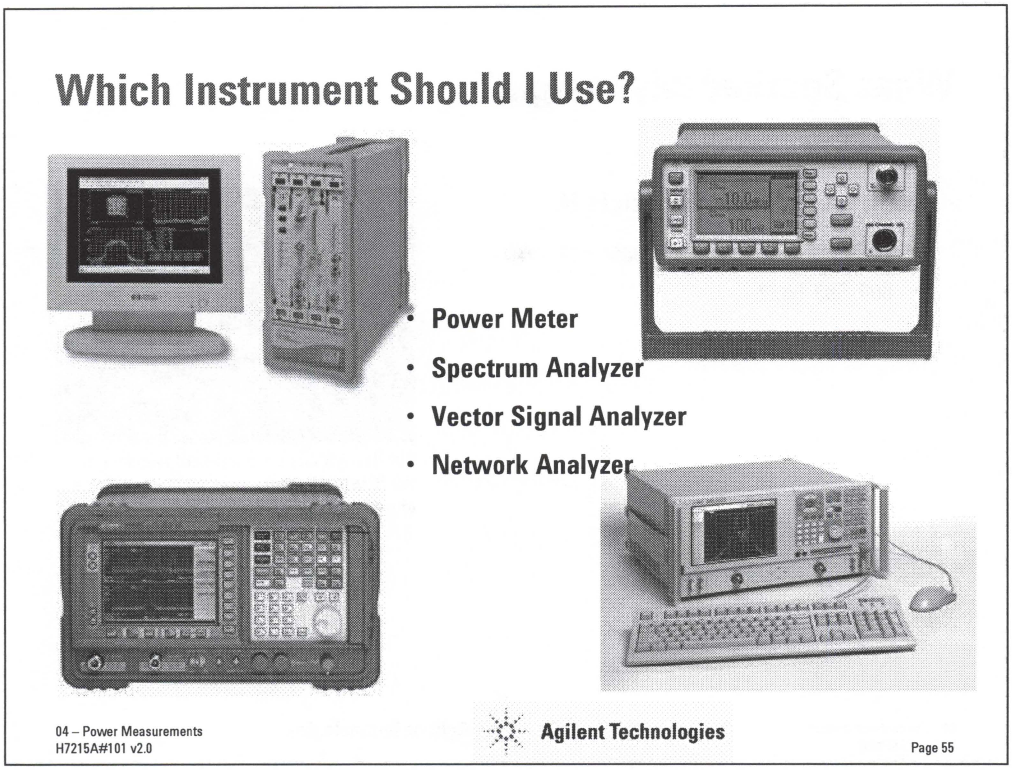
Case 1) is the measurement of the incident wave only (no reflection) and is how the ideal power sensor would behave.

Case 2) is a practical condition, if the power sensor (load) is nomimally Z_0 but not exactly Z_0 then the measurement is subject to both mismatch loss, and mismatch uncertainty. The *mismatch loss* is a quantity which may be determined so its effect is included in the calibration factor of the sensor. The *mismatch uncertainty* lies between limits determined from the magnitude of the source match and the load (sensor) match.

In all these cases the basis of the uncertainty has been Zo , the power *available* from the generator is not considered. Only the incident power on the load is considered.



We usually consider Z_0 to be real, but this is not necessarily so. The points noted above show that a reflectionless transfer is not always the maximum power transfer, and max power transfer is only reflectionless when the source and load impedances are real.



Although a variety of instruments measure power, the most accurate instrument is a power meter and a sensor. The sensor is an RF power-to-voltage transducer. The power meter displays the detected voltage as a value of power in log (dBm) or linear (watts) units. Typical accuracies of a power meter will be in the hundredths of a dB, while other instruments (i.e., spectrum analyzers, network analyzers) will have power measurement accuracies in the tenths of dBs or more.

One of the main differences between the instruments is that of frequency selective measurements. Frequency selective measurements attempt to determine the power within a specified bandwidth. The traditional Power Meter is not frequency selective in that it measures the average power over the full frequency range of the sensor and will include the power of the carrier as well as any harmonics which may be generated. A Spectrum Analyzer provides a frequency selective measurement since it measures in a particular Resolution Bandwidth.

The lack of frequency selectivity is the main reason that Power Meters measure down to around -90dBm and instruments such as a spectrum analyzer can measure much lower than this if very narrow resolution bandwidths are used.

What Sparked My Insight?

- **Spark your educational insight @**
  - www.agilent.com/find/tmeducation
  - **800-593-6632**

Bill Hewlett and Dave Packard founders of the Hewlett-Packard company and heritage of Agilent Technologies.

04 – Power Measurements
H7215A#101 v2.0

Agilent Technologies

Page 56

Appendix

How to Compare Things — Note on logs and dB's

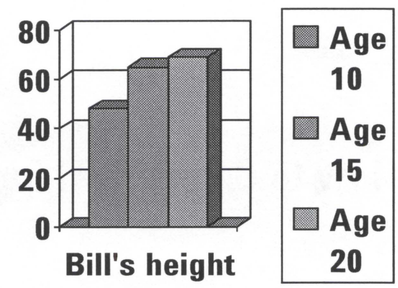
Agilent Technologies

In this lesson the decibel (dB) is explained:

- Comparing quantities using exponents or powers
- What is a logarithm
- The dB: a logarithmic power ratio
- The dBm: a way to express absolute power
- Using dBs and dBms together
- The dB as an expression of a voltage ratio
- An example of power and voltage ratios
- dBs used to express tolerances
- A table of power and voltage tolerances

Comparing Things

- In everyday experience there are two ways of expressing an increase or decrease.
  - As an addition or subtraction.
  - As a factor or multiple



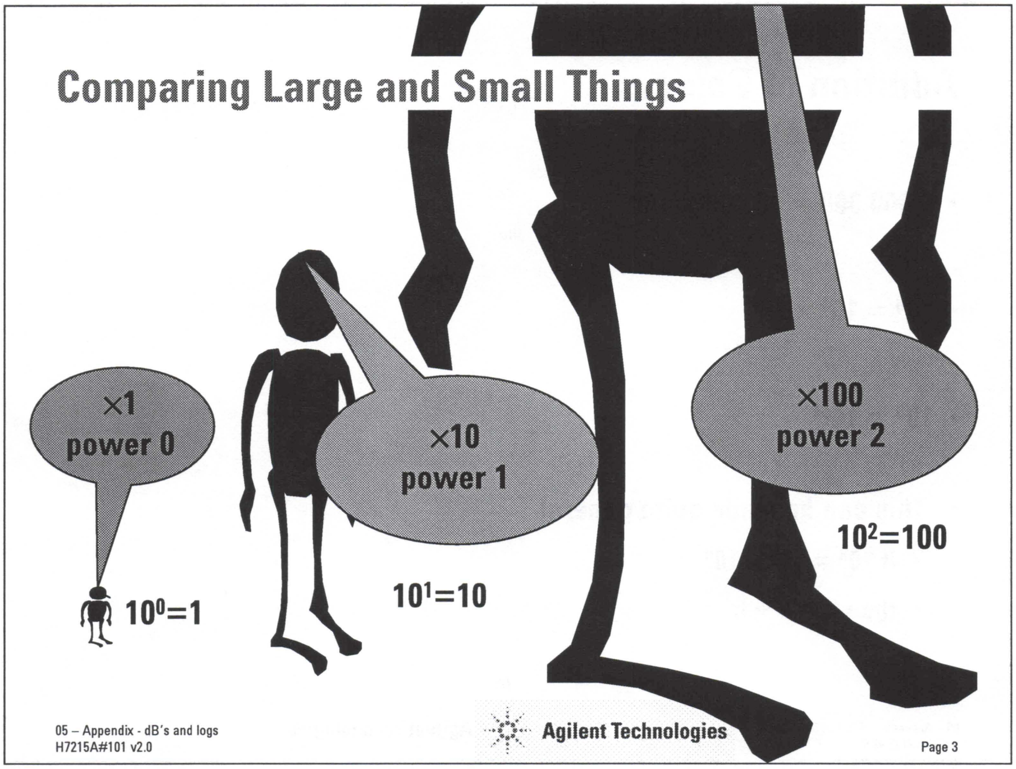
Bill's height increases by:
17" and 21" from 48" or ...
1.35 times. and 1.44 times his height at 10 years.

05 – Appendix - dB's and logs
H7215A#101 v2.0

Agilent Technologies

Page 2

There are different ways to compare quantities, here are some common examples.



Powers of Numbers

The way we express the power of a number is to give that number a superscript. If we choose the number ten and use whole numbers as the superscript the the meaning is clear. $10^2 = 10 \times 10$; $10^3 = 10 \times 10 \times 10$ etc.

The power of ten is the number of times ten has been multiplied by itself (which is also the number of zeros following the one).

Following the above rule it follows that $10^1 = 10$, and that $10^6 = 1{,}000{,}000$
The statement; ten multiplied by itself zero times makes no sense in our spoken language but in the logic of mathematics it follows directly from the addition property of powers.

In the above examples the number which has the assigned power is called the base; in this case the base is ten.

We use powers to *express very large or very small quantities in a compact way*. For example: velocity of electromagnetic waves in a vacuum (Radio and light waves) = 3×10^8 meters/sec

Addition of Powers

- $1{,}000{,}000 = 10{,}000 \times 100$
- $10^6 = 10^4 \times 10^2$
- $10^6 = 10^{(4+2)}$
- **This can be made quite general**
  - if $10^x = 10^a \times 10^b$
  - then $x = a + b$

05 – Appendix - dB's and logs
H7215A#101 v2.0

Agilent Technologies

Page 4

The addition of powers may be used to show that 10^0 has a meaning.
Rearranging the above:

$$10^b = 10^x \div 10^a$$

$b = x - a$; if $x = a$ then

$$b = 0 \text{ and } 10^b = 10^a \div 10^a = 1$$

You could deduce may results from this for example if $a > x$ then b is negative so $10^b = 10^x \div 10^a$ is less than one.

Some Results

- $10^6 = 1{,}000{,}000$
- $10^2 = 100$
- $10^1 = 10$
- $10^0 = 1$
- $10^{-1} = 1/10$
- $10^{-3} = 1/1000$
- $10^{0.3010} = ???$

What about fractional powers?

05 – Appendix - dB's and logs
H7215A#101 v2.0

Agilent Technologies

Page 5

So far we have given meaning to positive and negative powers expressed as whole numbers but what about fractional powers? It turns out that with some extension of the logic we have used we can find a meaning for fractional powers of any base. The base we shall be using for dB definition is base ten. With more advanced study other bases are used but need not concern us for this lesson. The answer to the above question is 2; that is $10^{0.3010} = 2$ to four significant figures.

Logarithms and dB

- Logarithms or Logs are a way of expressing multiples and factors in terms of a power index.
- For dB we use logs to the base ten. ($\log_{10}$)
  - For example $\log_{10}(100) = 2$ [$10^2 = 100$]
- Log gives the inverse of raising to a power.
- The number in parenthesis () is a power ratio
  - $dB = 10.\log_{10}(P_1/P_2)$
  - For example 20dB is equivalent to a power ratio of 100

05 – Appendix - dB's and logs
H7215A#101 v2.0

Agilent Technologies

Page 6

Logarithms are the inverse of a power expression just as division is the inverse of multiplication. In the previous example log(2) = 0.3010; notice the "10" subscript has been dropped, usually the word log without a base subscript is assumed to be to the base ten.

The decibel is based on the bel, named after Alexander Graham Bell.He was interested in with the way in which the human ear responds to sound intensity. He used a logarithmic scale to express this sound intensity, the range from the softest sound to the loudest (threshold of pain) sound, is one to a billion (10^{12}) or zero to 12 bels.

The decibel is one tenth of a bel.

We use dB to *express very large or very small ratios in a compact way.*

Some Useful dB's

Power Ratio	10logP	Usual dB	20logP	Usual dB
• 0.1	–10.00	–10	– 20.00	– 20
• 0.5	–3.01	–3	–6.02	–6
• 1.0	0.00	0	0.00	0
• 2.0	3.01	3	3.01	3
• 4.0	6.02	6	6.02	6
• 5.0	6.99	7	13.98	14
• 10.0	10.00	10	20.00	20
• 100.00	20.00	20	40.00	40
• NEG	not allowed; gives error on calculator			

05 – Appendix - dB's and logs
H7215A#101 v2.0

Agilent Technologies

Page 7

Many values may be derived from just two values.

1) 10log(2) = 3(.010) dB (the approximation 3 is usually used)

2) 10log(10) = 10dB

Power Ratio	**dB**	**Exact dB**(3 figs)
0.25	-6	
0.5	-3	
1	**0**	
1.25	1 (2.5 ÷ 2 is 4 – 3dB)	(0.969dB)
2	**3**	
2.5	4 (5 ÷ 2 is 7 – 3dB)	(3.97dB)
4	6 (twice 2 is 3 + 3dB)	
5	7 (10 ÷ 2 is 10 – 3dB)	
8	9 (twice 4 is 6 + 3dB)	(9.03dB)
10	**10**	
20	13	
50	17	
80	19	
100	**20**	

Lets Fill Them In ...

dB	0	1	2	3	4	5	6	7	8	9	10	
Ratio	1			2							10	

05 – Appendix - dB's and logs
H7215A#101 v2.0

Agilent Technologies

Page 8

The dBm

- The dB is an expression of relative power or a power ratio.
- The dBm is an expression of absolute power compared to 1mW, 10^{-3} W

$$\text{dBm} = 10 \log \left\{ \frac{P}{1\ \text{mW}} \right\}$$

05 – Appendix - dB's and logs
H7215A#101 v2.0

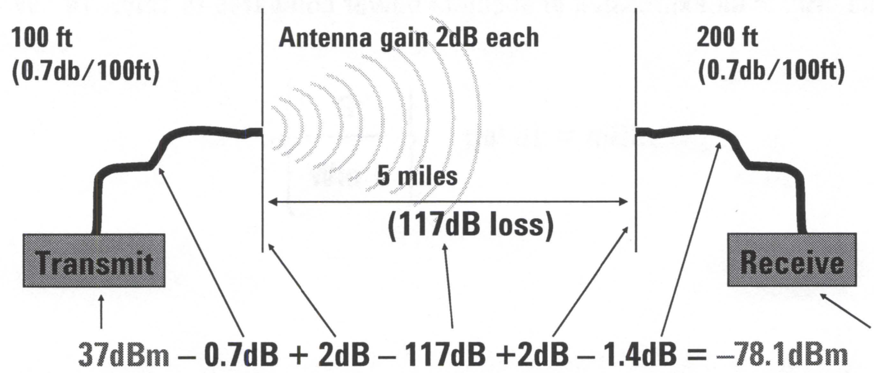
Agilent Technologies

Page 9

The dBm is the most common unit based on the dB, there are many more. For example dBW (absolute power compared to a watt). dBμW (compared to a microwatt)

A dB Example

- Absolute (dBm) and relative (dB) may be combined.
- Tx power = 5 watts at 2GHz. What is the received power?



05 – Appendix - dB's and logs
H7215A#101 v2.0

Agilent Technologies

Page 10

In this example we show how dB's may be added and subtracted to give the result in dBm. If the dB's are known its really quite a simple exercise. I leave it up to you to calculate this result without using dB's. -78.1 dBm is about 0.0155 μW

dB Voltage Ratios

- **For a common R**

$$\text{Power} = \frac{V^2}{R}$$

$$\text{Ratio expressed in dB} = 10\log\frac{P_2}{P_1} = 10\log\left(\frac{V_2}{V_1}\right)^2 = 20\log\left(\frac{V_2}{V_1}\right)$$

05 – Appendix - dB's and logs
H7215A#101 v2.0

Agilent Technologies

Page 11

In many lower frequency applications dB's are used to express a voltage ratio but the definition is still derived from a power ratio. This means that dbmV and dBµV are offset from dBm by a different amount depending on the impedance reference.

So $20\log(V_2/V_1)$ is the number of dB by which corresponding powers differ expressed in terms of voltage ratios.

dBm to dB(voltage) - Depends on Impedance

From dBm to ...

	dBμV	dBmV
50 Ω	dBm + 107dB	dBm +47dB
75 Ω	dBm + 108.75dB	dBm + 48.75dB

05 – Appendix - dB's and logs
H7215A#101 v2.0

Agilent Technologies

Page 12

$$20\log\frac{V}{1\text{mV}} = \text{POWER} = \text{dBmV}$$

$$V \text{ mV. will dissipate } \frac{V^2}{1000\boldsymbol{R}} \text{mW}$$

$$\begin{aligned} \text{dBm } &= 20\log V - 10\log(1000) - 10\log \boldsymbol{R} \\ &= \text{dBmV} - 30 - 10\log \boldsymbol{R} \\ \therefore \text{dBmV} &= \text{dBm} + 30 + 17 (\text{for } 50\Omega) \\ &= \text{dBm} + 47\text{dB} \\ &= \text{dBm} + 48.75\text{dB (for } 75\Omega) \end{aligned}$$

For dBμV add another 60dB

Absolute Voltage dB's

- dBmV is 20log(V/1mV)
- dBµV is 20log (V/1µV)

- Also field strength often expressed dBµV/meter

05 – Appendix - dB's and logs
H7215A#101 v2.0

Agilent Technologies

Page 13

To Get From dBm to dBμV

$$\text{dBW} = 10\log\frac{V^2}{R} = 20\log V - 10\log R$$

in dBm and microvolts

P_m power in mW
V_μ Voltage in μV

$$10\log\left(\frac{P_m}{1000}\right) = 20\log\left(\frac{V_\mu}{1000000}\right) - 10\log R$$

$$10\log P_m - 30 = 20\log V_\mu - 120 - 10\log R$$

$$\text{dBm} = 20\log V_\mu - 10\log R - 90$$

$$\text{dBm} = 20\log V_\mu - 17 - 90 \quad (\text{for } R = 50\Omega)$$

0dBm ⟶ 107dBμV (50 ohms) ⟶ 47dBmV (50 ohms)

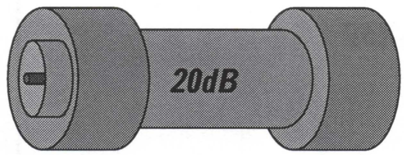
05 – Appendix - dB's and logs
H7215A#101 v2.0

Agilent Technologies

Page 14

Power and Voltage Ratio: Example

- **Apply 10 volts to the attenuator (Terminated in 50 ohms)**
- **How much Power is applied ?**
- **What is the output Power?**



V_{out} = ?

Z_0 =50 ohms

a) 1 volt
b) 0.1 volt
c) 10 volt

05 – Appendix - dB's and logs
H7215A#101 v2.0

Agilent Technologies

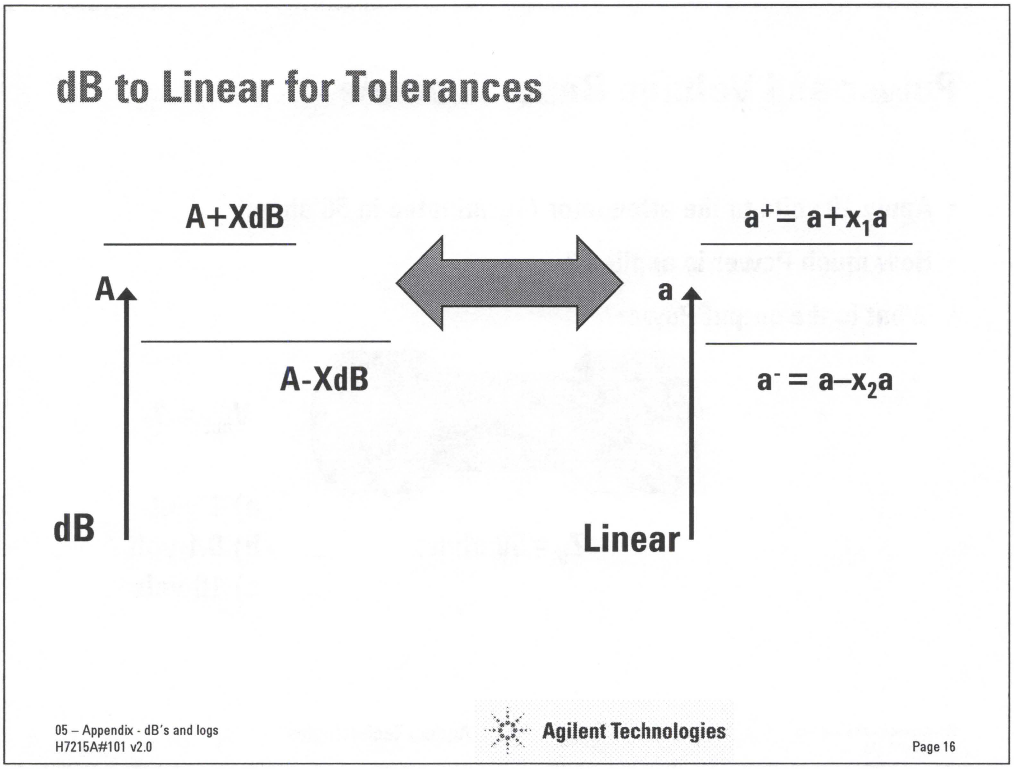
Page 15

Ten volts measured across 50ohms is 2 Watts into the attenuator *(which is the maximum average power specification for attenuators of the type 8491A)*

The voltage at the output of the attenuator would be 1 Volt into a 50 ohm termination which is 0.02 Watt; 20mW or 13dBm. This result could be arrived at by calculating (33 – 20)dBm.

Note: the dB value of the atenuator calculated for voltage ratio is the same as that calculated for the power ratio.

$$\text{dB value} = 20.\log(10) = 10.\log(100) = 20\text{dB}$$



In the following argument the logarithmic dB values are upper case (A)dB and the linear equivalent limits are lower case (a).

The dB tolerance values, (X), are added to (or subtracted from) the nominal value, (A), these limits are equivalent to the fractional increase and decrease on the linear nominal value, (a).

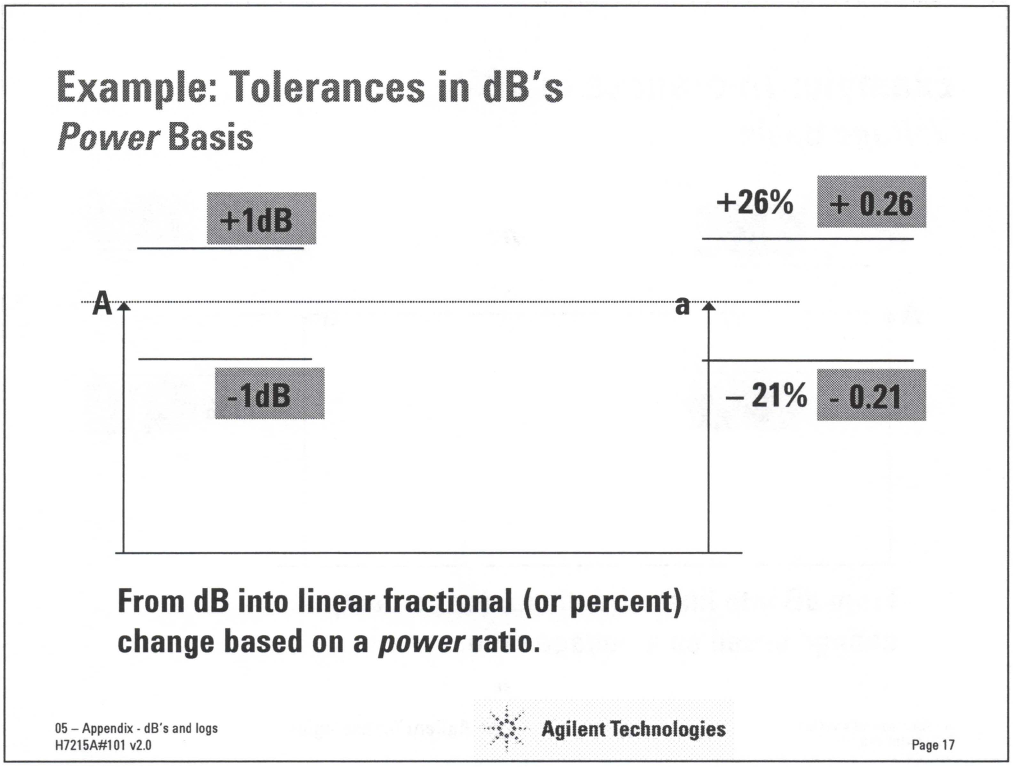
A maximum limit corresponding to (A + X)dB is $a(x_1 + 1)$ which is (a) the nominal value plus $(a.x_1)$ the fractional increase.

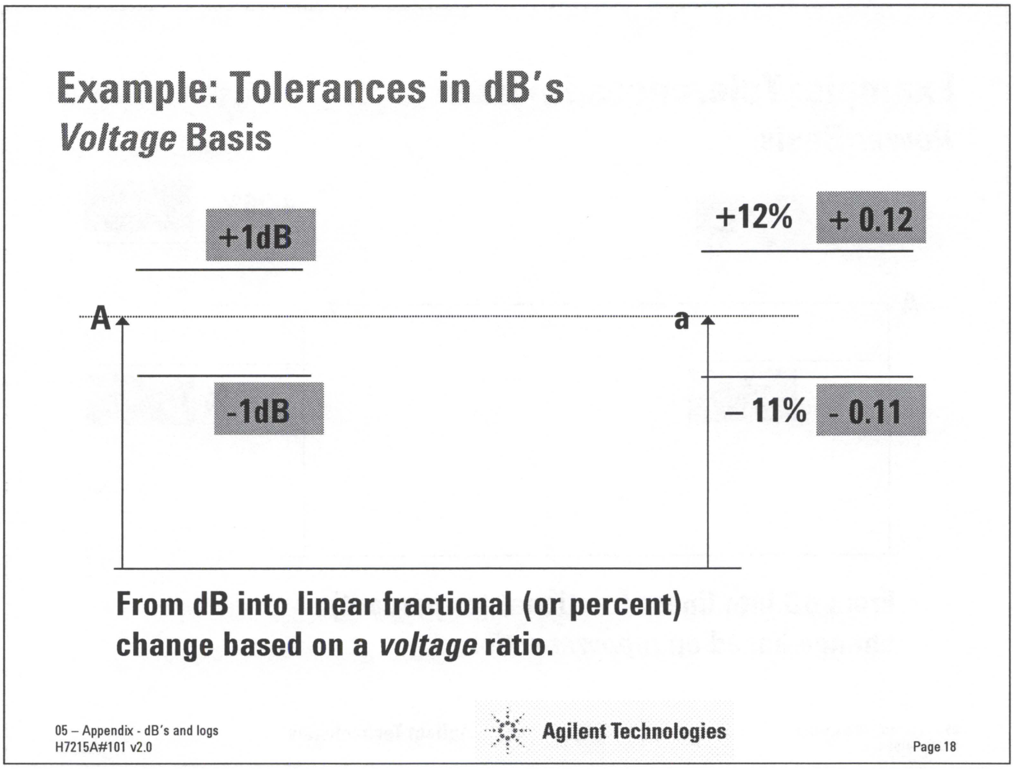
$$(A+X)dB = 10\log(a) + 10\log(1+x_1)$$
$$so\ \mathrm{X} = 10\log(1+\mathrm{x}_1)$$

$$(A-X)dB = 10\log(a) + 10\log(1-x_2)$$
$$so\ (-X) = 10\log(1-x_2)$$

For example if X = 1dB then the fractional increase is 0.26 and the fractional decrease is 0.21. (Based on a *power* ratio)

$(x_1 + 1) = 10^{0.1} = 1.2589$ and $(1 - x_2) = 10^{-0.1} = 0.7943$


Example: Tolerances in dB's
Power Basis
+1dB
-1dB
A
a
+26%
+ 0.26
– 21%
- 0.21
From dB into linear fractional (or percent) change based on a *power* ratio.
05 – Appendix - dB's and logs
H7215A#101 v2.0
Agilent Technologies
Page 17



If the calculation was based on a voltage ratio, then we calculate $10^{dB/20}$

For example if X = 1dB then the fractional increase is 0.22 and the fractional decrease is 0.11. (Based on a *voltage* ratio)

$(x_1 + 1) = 10^{0.05} = 1.1220$ and $(1 - x_2) = 10^{-0.05} = 0.8913$

For small dB values (<1dB) the fractional increase and decrease for a voltage ratio is approximately half that of the corresponding power ratio.

Percent Tolerance Table

dB to Percent Tolerance Table for Power and Voltage				
dB	Plus Pwr	Minus Pwr	Plus V	Minus V
0.01	0.23	0.23	0.12	0.12
0.05	1.16	1.14	0.58	0.57
0.1	2.33	2.28	1.16	1.14
0.5	12.20	10.87	5.93	5.59
0.8	20.23	16.82	9.65	8.80
1	25.89	20.57	12.20	10.87
2	58.49	36.90	25.89	20.57
3	99.53	49.88	41.25	29.21
5	216.23	68.38	77.83	43.77
10	900.00	90.00	216.23	68.38

Above line the voltage % is approximately half the power %

05 – Appendix - dB's and logs
H7215A#101 v2.0

Agilent Technologies

Page 19

Prefixes

1 000 000 000 000	$= 10^{12}$	tera T
1 000 000 000	$= 10^{9}$	giga G
1 000 000	$= 10^{6}$	mega M
1 000	$= 10^{3}$	kilo k
0.001	$= 10^{-3}$	milli m
0.000 001	$= 10^{-6}$	micro μ
0.000 000 001	$= 10^{-9}$	nano n
0.000 000 000 001	$= 10^{-12}$	pico p
0.000 000 000 000 001	$= 10^{-15}$	femto f

05 – Appendix - dB's and logs
H7215A#101 v2.0

Agilent Technologies

Page 20

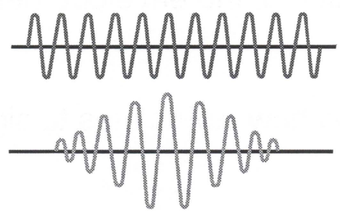
Prefixes

These are most of the ISO unit prefixes, the pronunciation of giga is usually with a hard g as in "get". I have sometimes heard a soft pronunciation as in "jet".

The symbol for micro (m) is sometimes shown as (u) on instruments and texts that cannot support Greek characters.

RF & Microwave Measurement Fundamentals

Amplitude Modulation



Agilent Technologies

Objectives

Explain why carrier modulation is used when Transmitting information

Introduce AM

Describe the term "AM modulation Index"

Describe the operation of the envelope detector

AM Measurements

Single Sideband and how AM relates to signal multiplexing

Agilent Technologies

SLIDE 2

Why do We Modulate?

What is modulation and why is it done?

Agilent Technologies

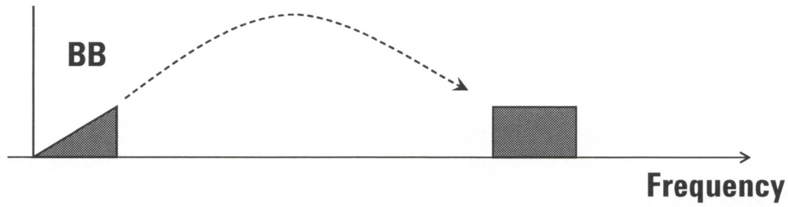
SLIDE 3

In radio frequency communications modulation is the impressing of an information signal on to a high frequency carrier.

What is Modulation

Process to encode information suitable to transmit.

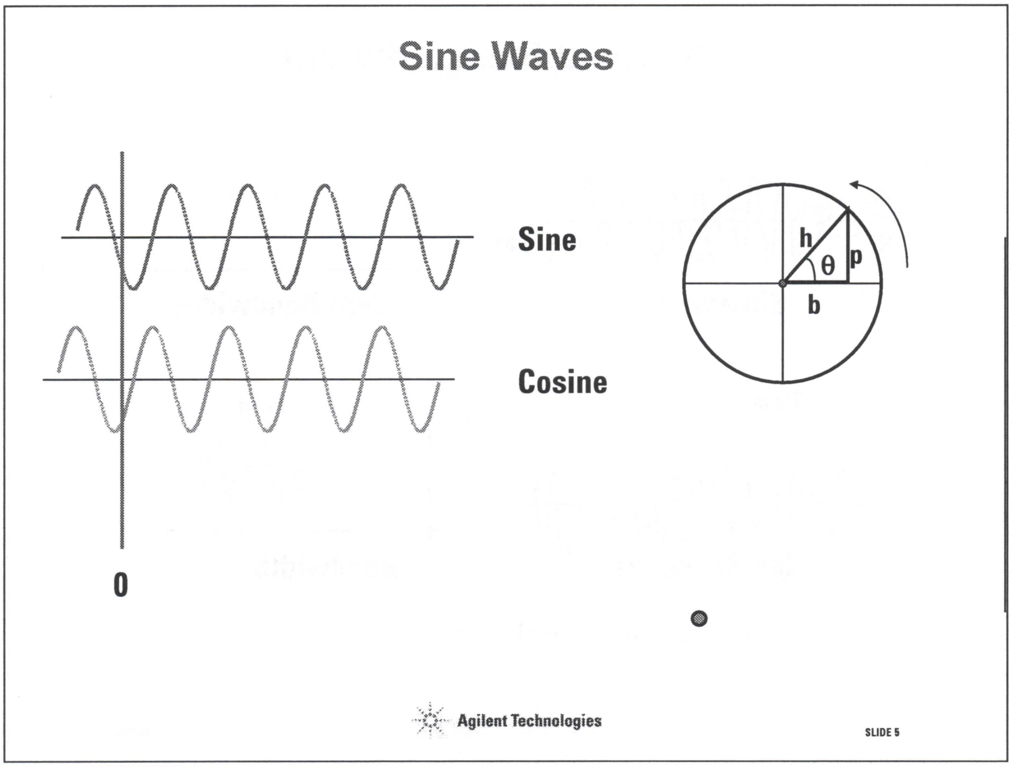
But usually means translating the baseband signal to a much higher frequency bandpass signal.



Agilent Technologies

SLIDE 4

The information signal, speech, video, data etc. is referred to as a *baseband* signal. This term can also mean the information signal to be modulated, this signal may not be the original information signal but the result of previous modulations or multiplexing.



Sinewaves in Trigonometry

A sinewave is the basis for all signals, so it is important to spend some time on the subject of sinewaves. In trigonometry the ratio of the sides of a right triangle are the sines and cosines. If the bottom or base is (b) units and the vertical or perpendicular is (p) units, and the remaining side, called the hypotenuse is (h) units then:

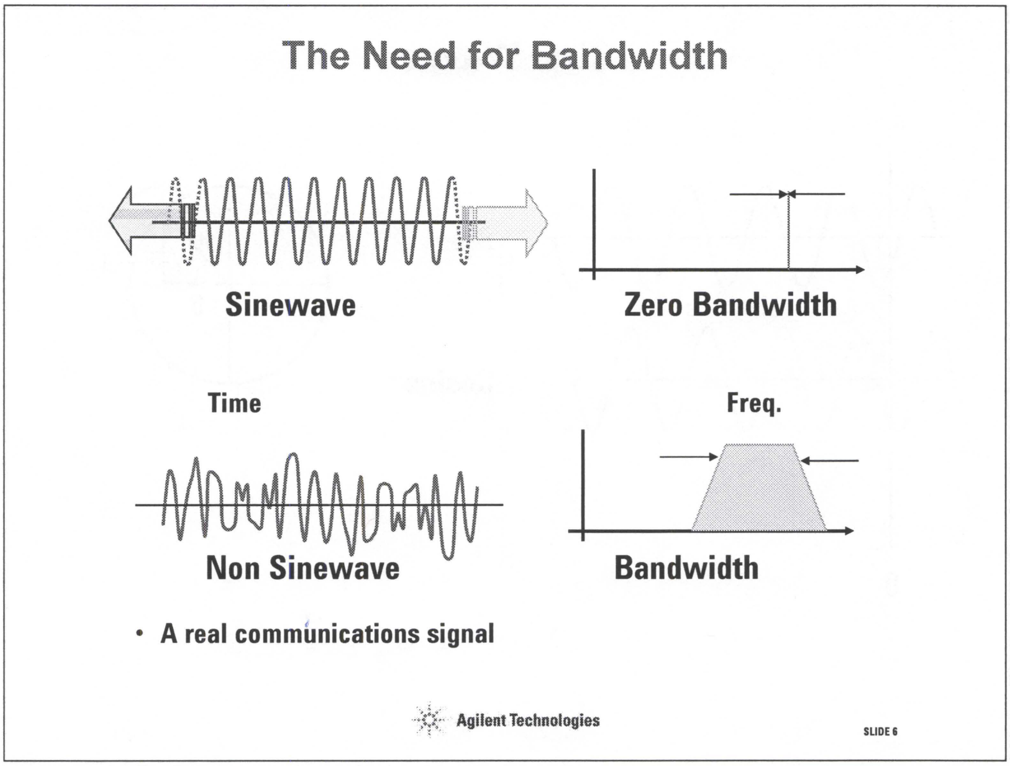
- $\sin \theta = p/h$ and $\cos.\theta = b/h$

These ratios are used to great effect by surveyors but are probably more useful in communications. If $h = 1$ then as h rotates about the axis, the angle θ is increased, from zero, then a graph of p would be a graph of the sine, and a graph of b the cosine.

Notice that the sine and cosine graphs are the same shape but have different starting points.

Why is this important?

The sine wave is the fundamental building block for our understanding of more complicated signals.

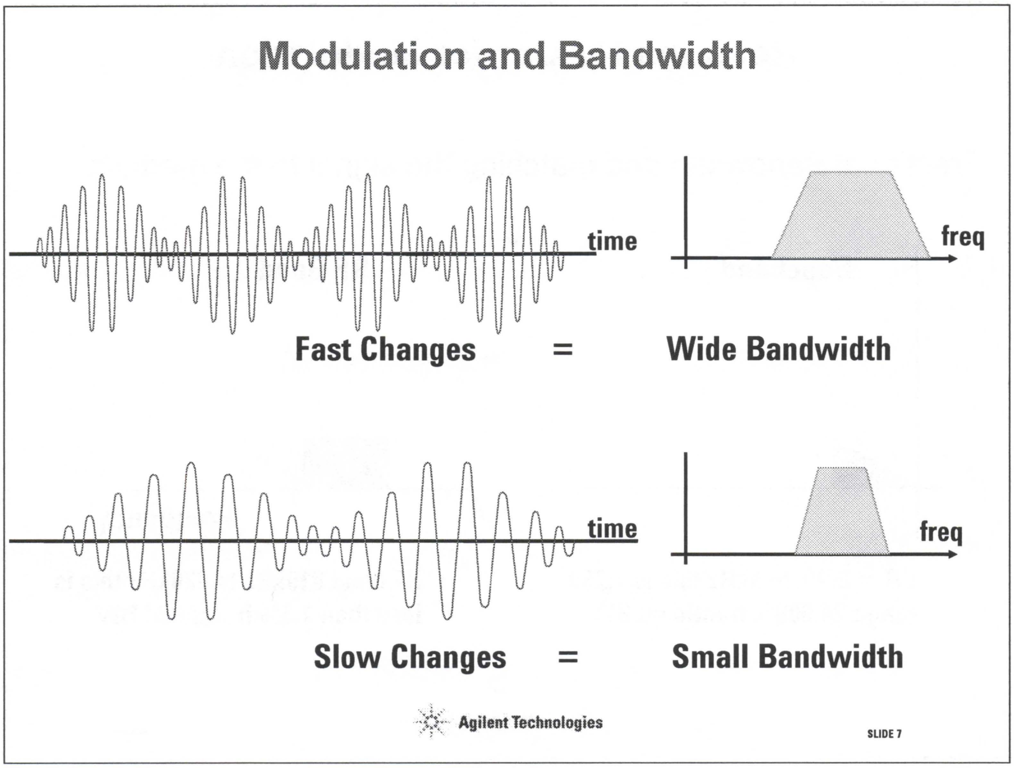


A sinewave is a very useful signal but unless something is done to it (modulation) no information can be transmitted.

The picture indicates that a perfect sinewave has existed and will exist forever. If anything is done to this basic wave such as switching it on and off (pulse modulation)* or varying the phase (phase modulation) or any other change the spectral component will either spread or require extra spectral lines or both.

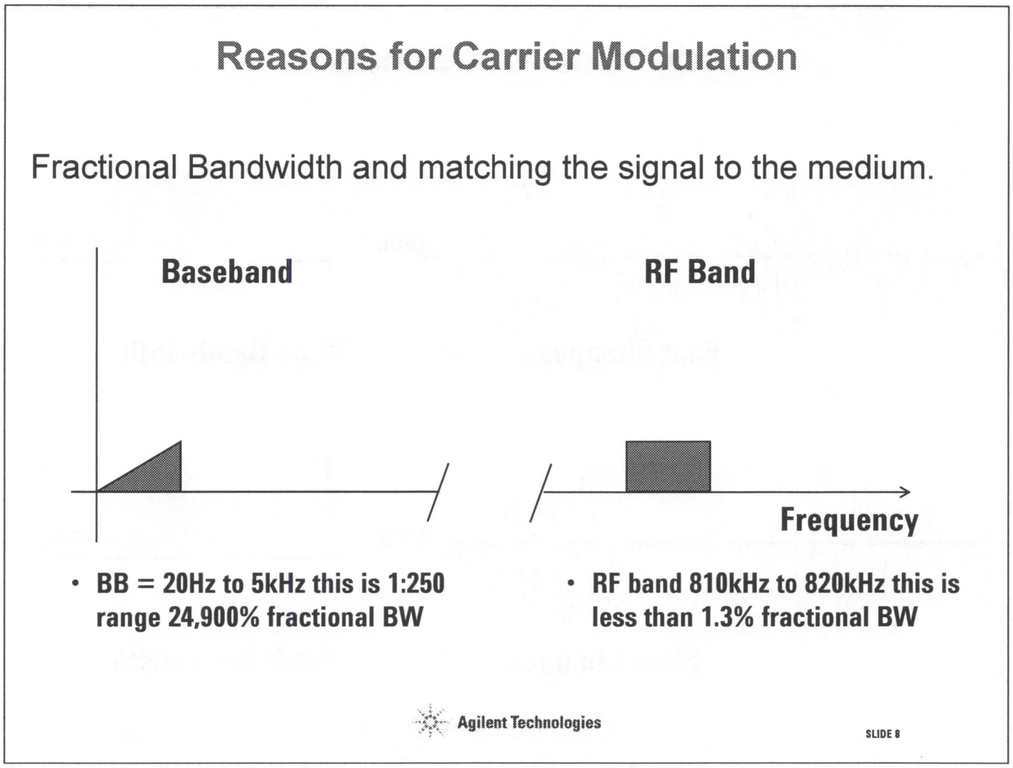
In a communications system there must be enough bandwidth to pass an acceptable representation of the signal. As we have seen limiting the bandwidth of the squarewave changed it's shape.

- * Even the normal switching on and off of a sinwave source will produce some spreading of the spectral line, but vanishing small and unmeasurable.



The carrier frequency chosen will center the band containing the information. This band may be wide or narrow depending on the rate of change of the parameters of the carrier by the baseband signal.

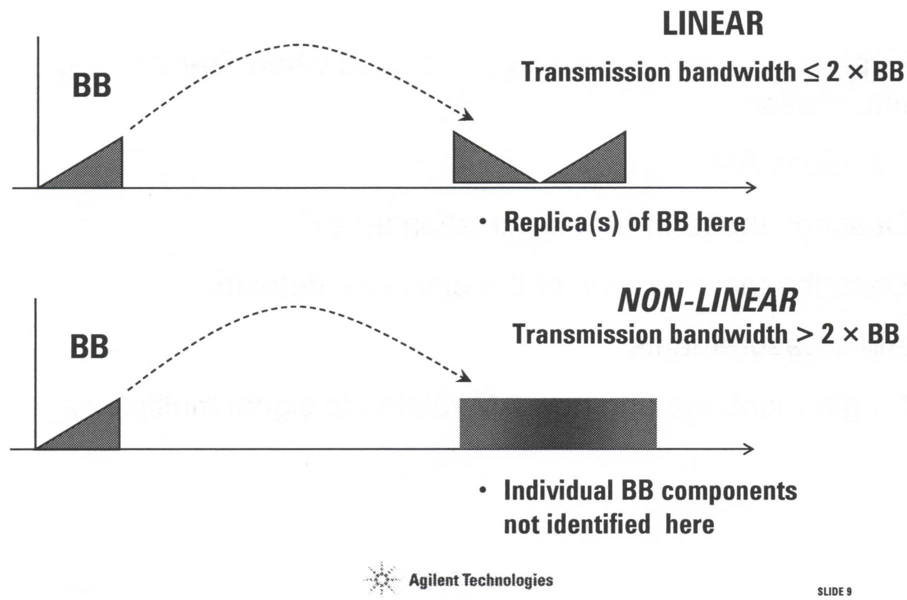
This use of a carrier frequency also allows a very important feature; the ability to *multiplex* several independent channels at different carrier frequencies. The occupied bands of each channel must not overlap.



This principle appears often in RF design. Here it is for the very practical consideration of antenna size. An antenna operating at the baseband frequency would be considered a very broadband device, very impractical for size and construction. The high frequency carrier allows an antenna of manageable size and a design only to handle a narrow band of frequencies.

The term "baseband" BB is the required signal band.

Linear and Non-Linear Modulation



Objectives

Explain why carrier modulation is used when Transmitting information

Introduce AM

Describe the term "AM modulation Index"

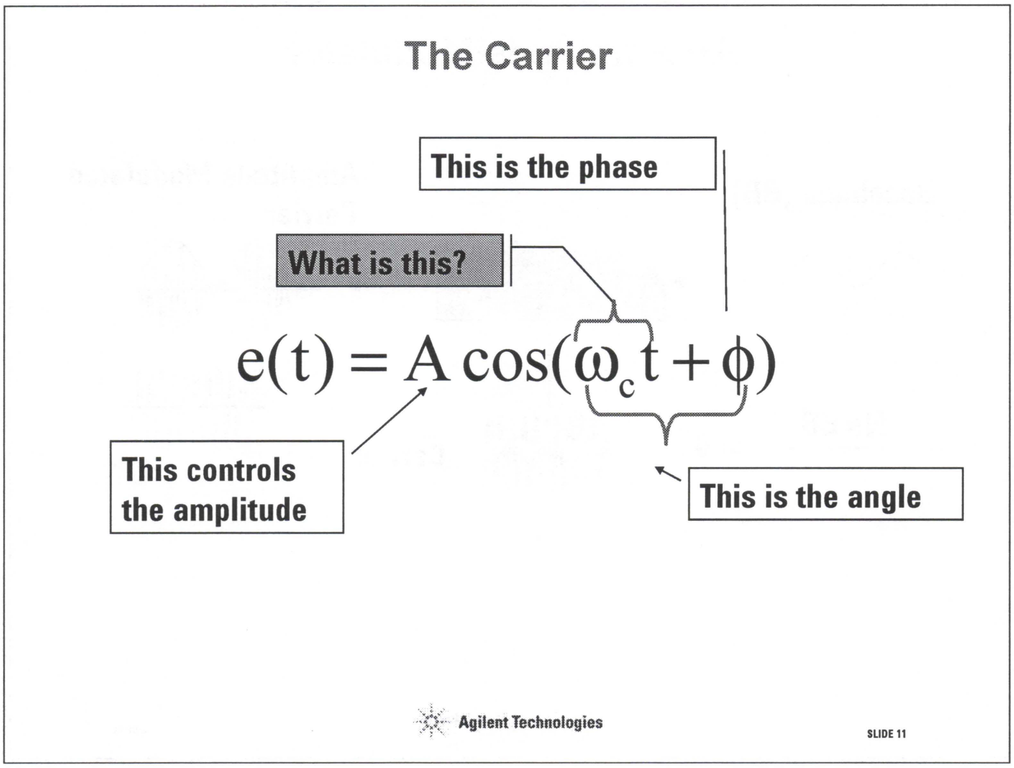
Describe the operation of the envelope detector

AM Measurements

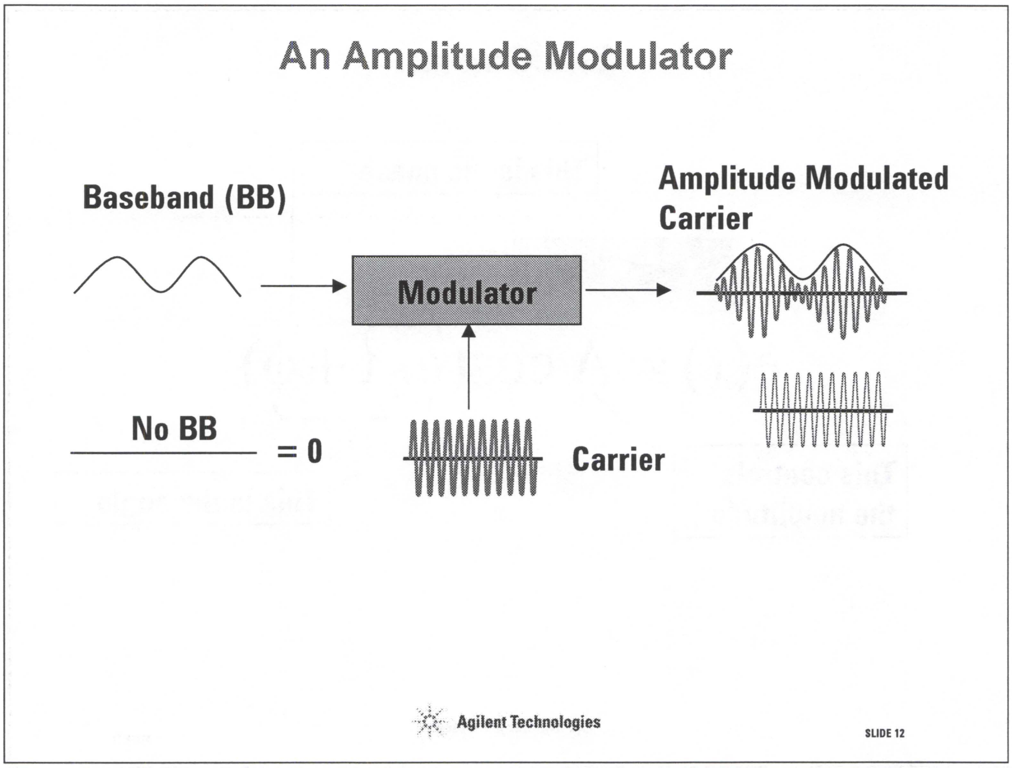
Single Sideband and how AM relates to signal multiplexing

Agilent Technologies

SLIDE 10



In amplitude modulation it is the amplitude (A), that is modulated, but it is important to understand the other parts of the "sinewave" function. There are two parts to the angle, the ϕ which may be described as the angle when $t = 0$ and $\omega_c t$ which is an ever increasing angle with time (angular frequency (ω_c) times time (t) = an angle).



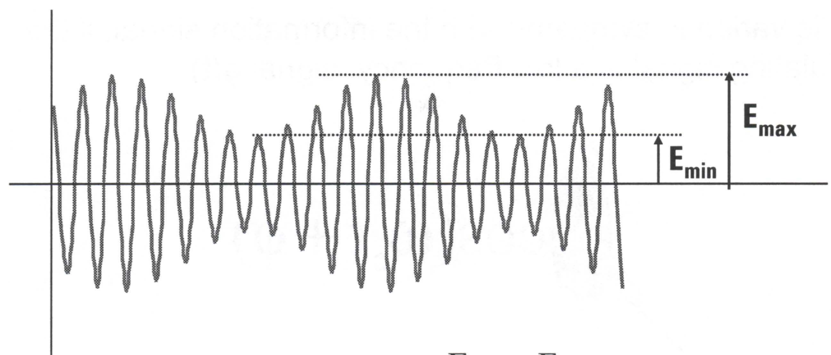
Without considering the formula, much information may be gathered informally by making some assessment of this picture.

If the BB signal is off, = 0, then the carrier is still present at its unmodulated level.

If the modulating signal (BB) is too strong then the negative peaks of the BB signal will try to drive the carrier envelope past its axis. So the peak BB amplitude ≤ the peak carrier. When the BB peak = carrier peak then the modulation index is unity or 100%.

The total power in the modulated signal is at least the power of the carrier, as the strength of the modulating signal is increased it appears that the total power of the modulated signal also increases. This last point will be shown analytically later.

Modulation Index



$$\text{modulation index} = \frac{E_{max} - E_{min}}{E_{max} + E_{min}} = m$$

Agilent Technologies

SLIDE 13

The modulation index describes the degree of modulation, in AM transmission m is constrained to be less than 1. If m>1 then the envelope would cross the 0 axis, such a signal would be distorted.

The modulation index is usually expressed as a percentage, so m = 1 is 100% modulation.

Linear Modulation

This is varied in sympathy with the information signal; if the modulating signal is a low frequency signal g(t).

$$A\cos(\omega_c t + \phi)$$

$$[1 + g(t)]\cos(\omega_c t + \phi)$$

Agilent Technologies

SLIDE 14

In amplitude modulation it is only (A) that is changed, but as we have just decided the carrier is still present when the modulation is zero. This means that (A) is only zero if the modulating signal is strong enough to make the envelope just touch the axis at its most negative point. So the modulating function g(t), is such that g(t) is ≥ -1. For all (t).

Objectives

Explain why carrier modulation is used when Transmitting information

Introduce AM

Describe the term "AM modulation Index"

Describe the operation of the envelope detector

AM Measurements

SIngle Sideband and how AM relates to signal multiplexing

Agilent Technologies

SLIDE 15

AM with a Sinusoidal Modulating Frequency

$$e(t) = [1 + m.\cos(\omega_m t)]\cos(\omega_c t)$$

$= \cos(\omega_c t) +$	• **Carrier plus ...**
$\frac{m}{2}\cos(\omega_c t - \omega_m t) +$	• **Lower sideband plus ...**
$\frac{m}{2}\cos(\omega_c t + \omega_m t)$	• **Upper sideband.**

Agilent Technologies

SLIDE 16

AM with a sinusoidal modulating frequency.

By replacing the general baseband function g(t) with $\cos(\omega_m t)$
From the trigonometry identity:

- **cosA.cosB = ½(cos(A – B) + ½(cos(A + B)**

The resulting AM signal has three parts:

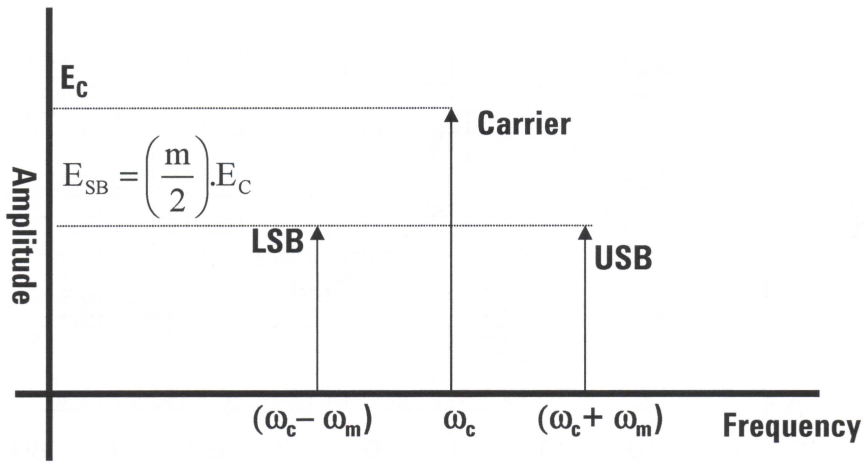
- Carrier,
- Lower sideband
- Upper sideband.

In this analysis the two sidebands are weighted by m/2 meaning that the upper and lower sidebands are 1/2 the voltage of the carrier if m = 1.

m is the modulation index, and is usually expressed as a percentage,

- $m\% = m \times 100$

AM in the Frequency Domain



Agilent Technologies

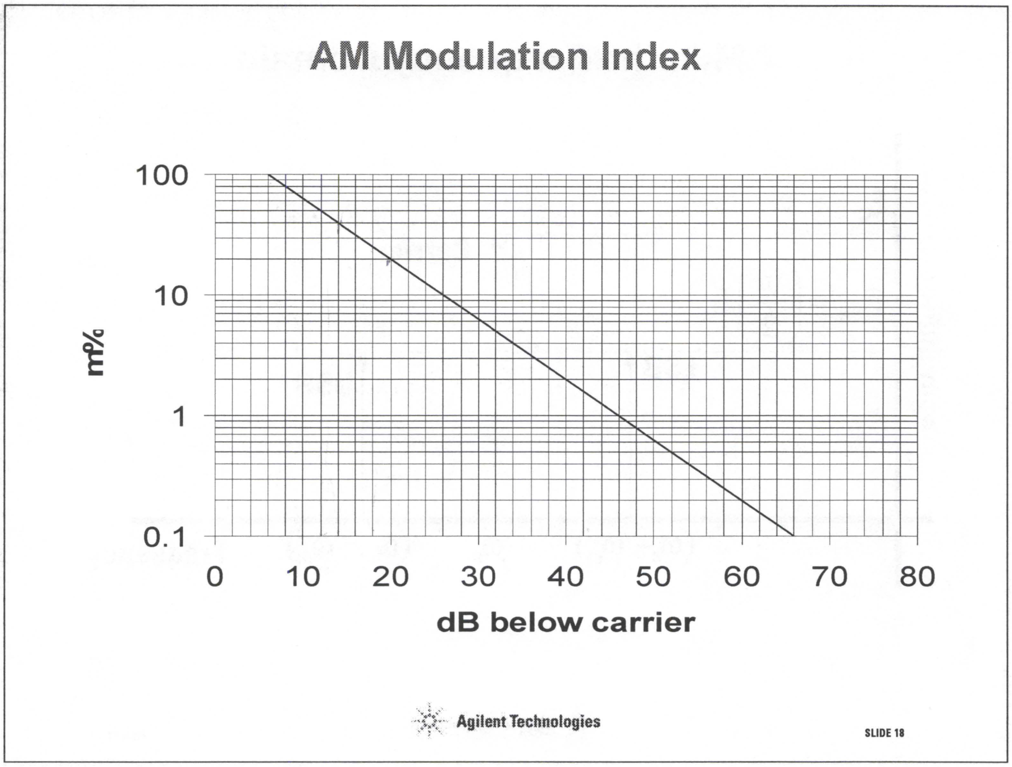
SLIDE 17

The modulation index may be expressed in terms of the sideband voltages, so that: $m = 2E_{SB}/E_C$. When using a spectrum analyzer to measure modulation index it is useful to have this equation in dB form:

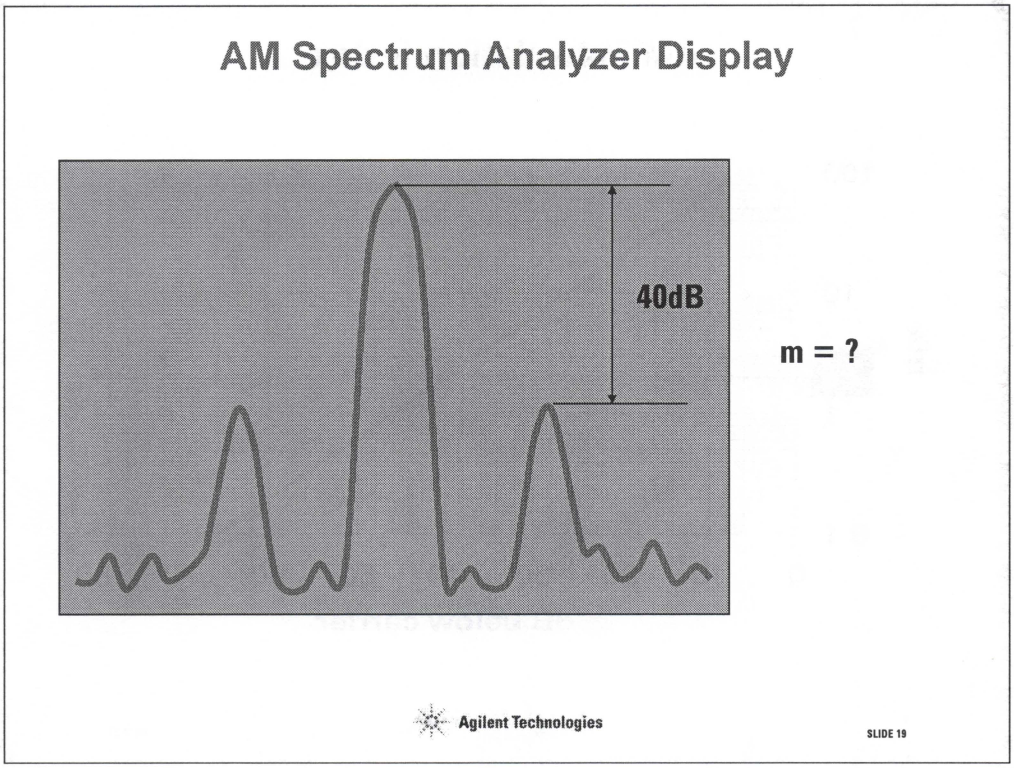
- (sideband difference in dB) = 20log(m/2)
- = 20log(m) – 6
- m% = [10^(6 – |sideband difference dB|)/20] × 100

This is quite easy: take the difference in dB (a positive number)

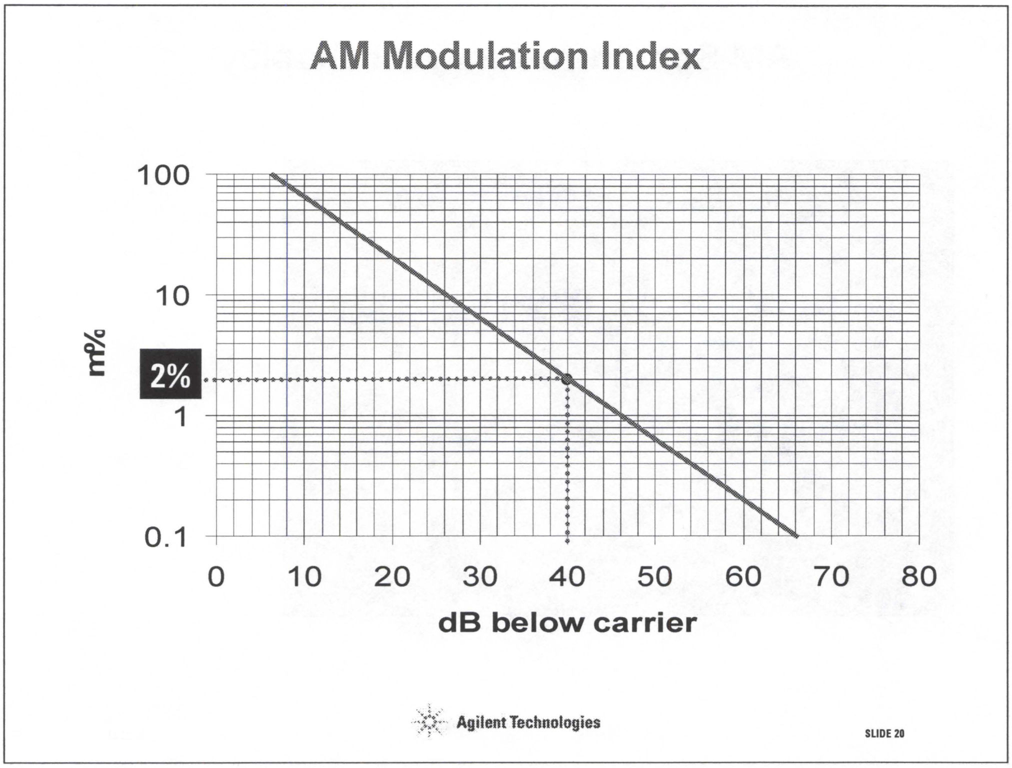
- Subtract 6
- Change sign (result is now a negative number) = x
- Divide by 20
- 10^x times 100



This graph will help to find the modulation index from the dB difference between the sidebands and the carrier.

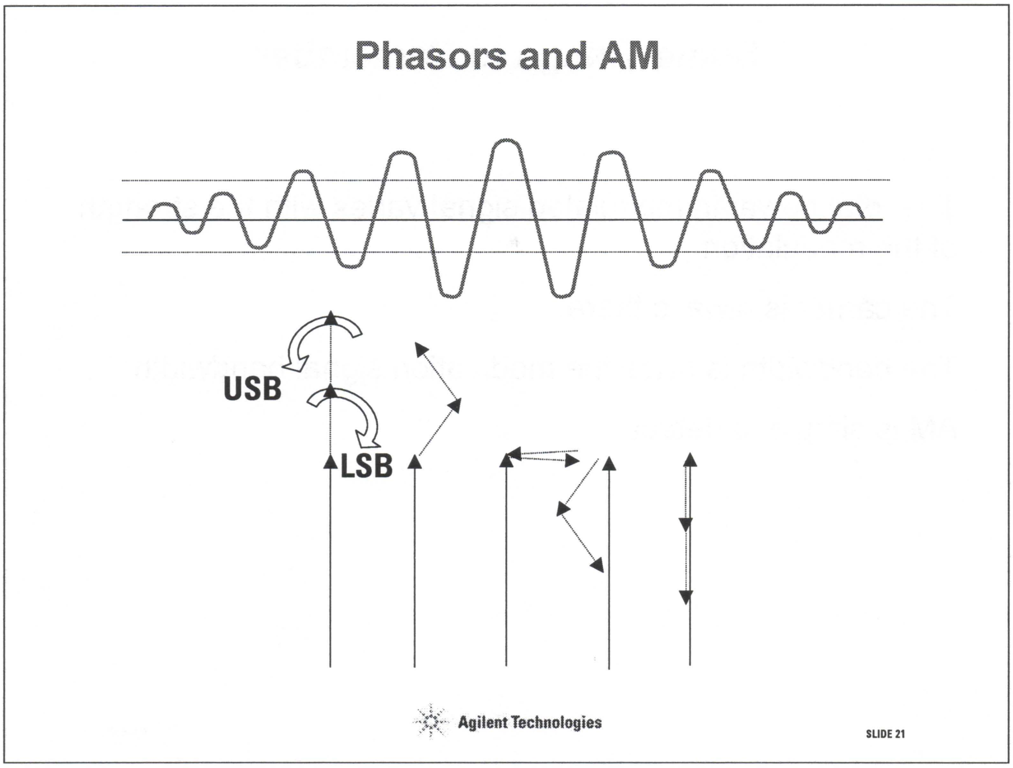


Use the formula and check against the graph to measure the modulation index for the AM signal shown here.



To calculate:

- Take the difference in dB (a positive number) 40
- Subtract 6 34
- Change sign (result is now a negative number) = x -34
- Divide by 20 -1.7
- 10^x times 100 2%



The carrier and sidebands may be represented by phasors. The carrier phasor is stationary in the above illustration. This is because we are only interested in their relative motion, so we as observers rotate at the same speed as the carrier phasor. The result of adding the phasors is that the carrier phasor periodically gets longer an shorter.

Note that the phasor does not rock from side to side because the USB and LSB are perfectly balanced. Imagine the effect of applying an amplitude modulated signal to a device which can unbalance them, the result would be this rocking side to side. This effect is called AM to PM conversion.

In actual measurements, sidbands are sometimes not equal in size, AM to PM conversion in the device being measured may account for this.

Some Things to Remember

The total power in modulated signal varies with the strength of the modulation

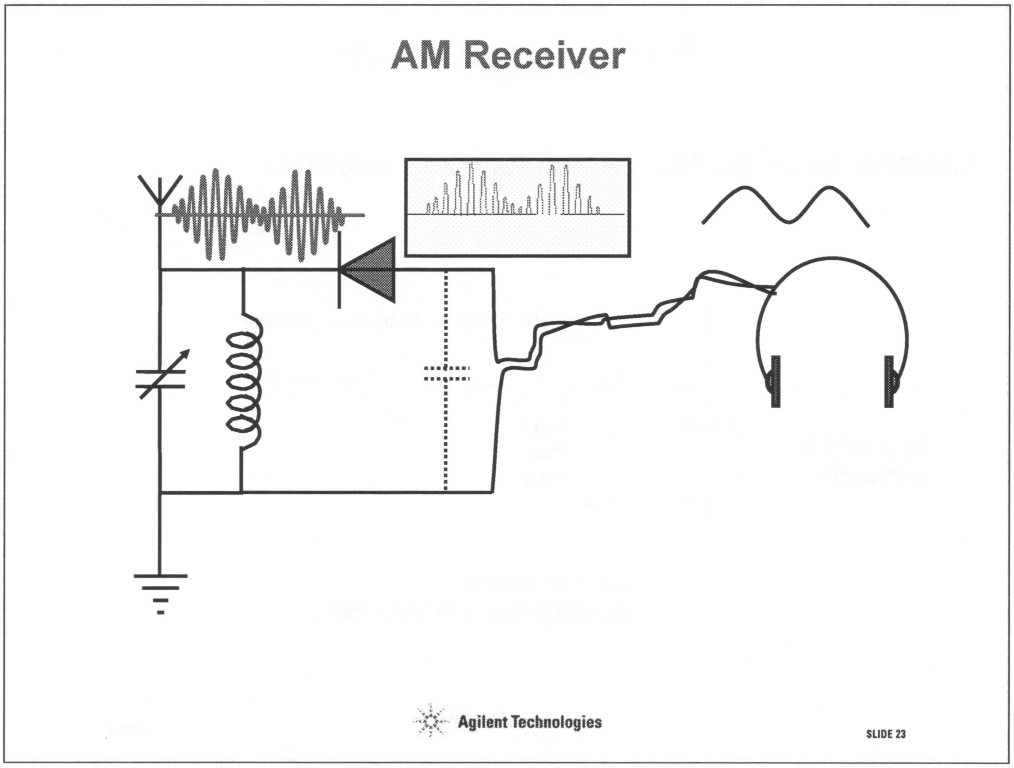
The carrier is always there

The bandwidth is twice the modulation signal bandwidth

AM is simple to detect

Agilent Technologies

SLIDE 22

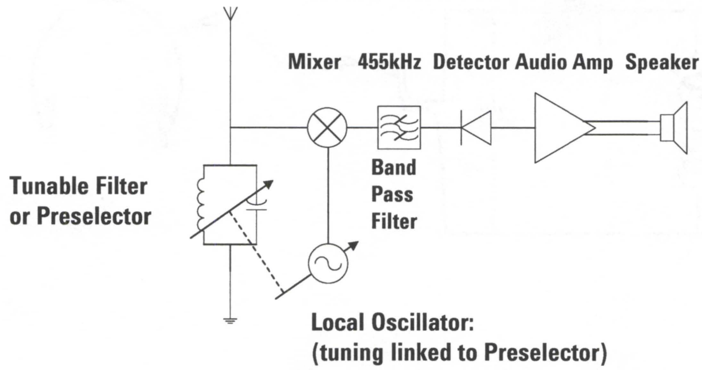


The basis for AM receivers "The Crystal Set"

Here is the most basic AM receivers. The tuned circuit is adjusted to resonate at the carrier frequency, the diode rectifies the signal. Usually there is enough capacitance in the connection to the headphones to loose (decouple) the high frequency, in any case the headphones cant respond to the high frequency.

The Superhet Receiver

Antenna receives AM band 540kHz to 1600kHz



Agilent Technologies

SLIDE 24

This receiver was the breakthrough in radio design, invented by Armstrong and first published in 1921, this is the basis of all modern receivers. In this case we show a simple AM radio set. Notice that the input is pre-selected to ensure that only one image of the RF spectrum is downconverted. The question is should we select the "high" or "low" side. Since:

$F_{IF} = F_{RF} - F_{LO}$ Low side LO,

$F_{LO} = F_{RF} - F_{IF}$

or $F_{IF} = F_{LO} - F_{RF}$ High side LO

$F_{LO} = F_{IF} + F_{RF}$

This is where we must consider designing an LO to cover enough range to tune the radio over the required RF spectrum. In the first case the LO would have to tune from a very low frequency (85kHz) to 1145kHz a range of 13:1.In the second case the LO would tune from 955kHz to 2055kHz a range of 2:1.

Which oscillator would be easiest to implement?

Objectives

Explain why carrier modulation is used when Transmitting information

Introduce AM

Describe the term “AM modulation Index”

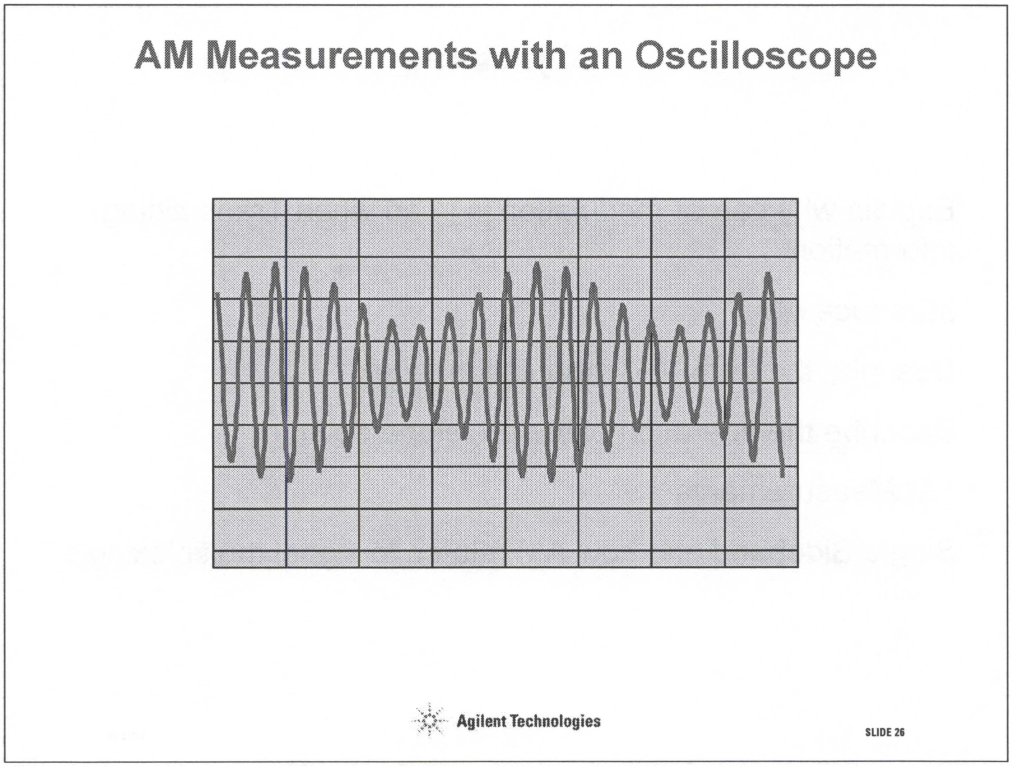
Describe the operation of the envelope detector

AM Measurements

Single Sideband and how AM relates to signal multiplexing

Agilent Technologies

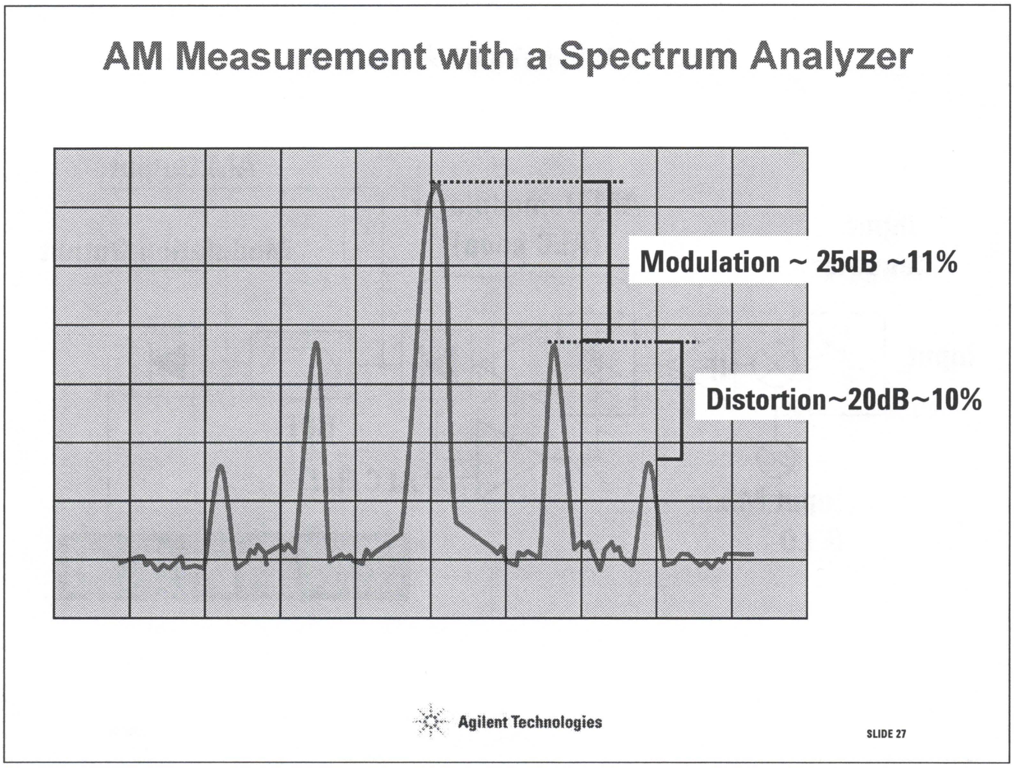
SLIDE 25



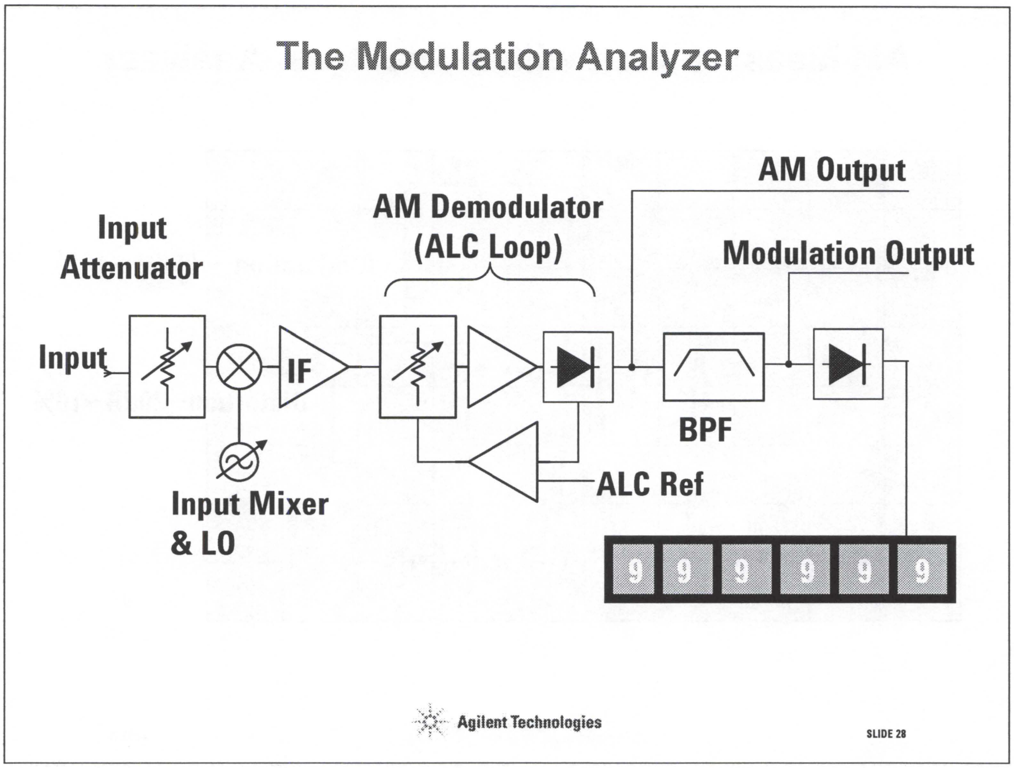
An oscilloscope display may be used to measure AM modulation index for sinewave baseband signal.

$$\text{modulation index} = \frac{E_{max} - E_{min}}{E_{max} + E_{min}} = m$$

Accuracy is limited,depending on the amount of modulation, to between 1.5% and 10%. The display resolution, linearity and frequency response of the oscilloscope will limit the usefulness of this technique.



A spectrum analyzer can show very small modulations, as low as 0.1%, with a sinusoidal modulating signal. With real modulating signals such as voice or music the spectrum analyzer is still useful in estimating the maximum modulating signal.



An example of an instrument dedicated to modulation measurement is shown here. This diagram is based on the HP8901B Modulation Analyzer and is a special receiver with an accurate detector in the IF (Intermediate Frequency) section.

Objectives

Explain why carrier modulation is used when Transmitting information

Introduce AM

Describe the term “AM modulation Index”

Describe the operation of the envelope detector

AM Measurements

Single Sideband and how AM relates to signal multiplexing

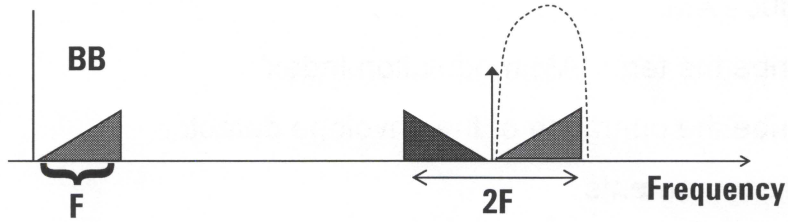
Agilent Technologies

SLIDE 29

Single Sideband

All the information is in one of the sidebands

Why not transmit just one of them?



Agilent Technologies

SLIDE 30

It occurred to engineers very early in the days of AM that a lot of power could be saved and double the bandwidth efficiency by transmitting just one of the sidebands.

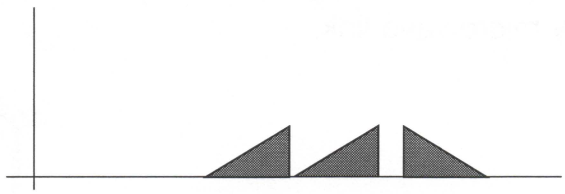
In the example above we have selected just one of the sidebands, the upper one is in the same direction as the BB signal so there is less processing to do. Most SSB transmitters don't use the filter as shown but use a special modulator similar in concept to the IQ modulators used in digital modulation. It's interesting to note that if the lower sideband were used without making it erect the sound would be recognizable as speech but that's all.

The operation of the SSB receiver is much more complicated than a simple envelope detector, and involves reinserting the carrier.

SSBSC Single Sideband Suppressed Carrier communications are used by radio amateurs, police, fire, and other mobile services.

Formation of Multiplex with Single Sideband AM

Single Sideband signals stacked in the frequency domain to create a multi-channel system. Carriers are suppressed.



Agilent Technologies

SLIDE 31

The property of modulating on to a carrier allowed channelized communications , and this is called *multiplexing*.

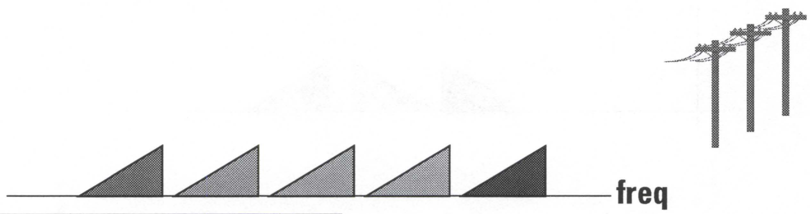
When many communication channels must use the same medium the multiplexing must be as efficient as possible ; SSBSC (Single Sideband Suppressed Carrier) allows very efficient packing of channels.. If single sideband channels are each given a different modulation frequency then they may be arranged to occupy different bands in frequency spectrum of the transmission system, as shown. Since we are dividing up the frequency spectrum we call this *Frequency Division Multiplex or FDM*.

An FDM may have a mixture of erect and inverted sidebands, the layout of the modulation plan was done in order to minimize intermodulation between channels. This is related to the "adjacent channel power" or ACP specification in a cellular system.

Where is SSB Used?

This is a Frequency Division Multiplex (FDM) as used in wireline telephony.

As many as 1800 4kHz voice channels may be stacked for transmission by microwave link.



A very important legacy system.
Introduce the idea of multiplex.

Agilent Technologies

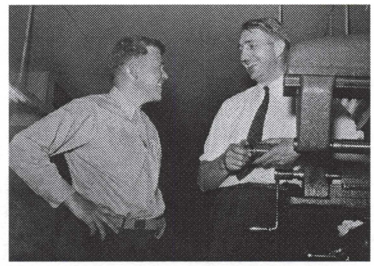
SLIDE 32

A very significant application was in wireline telephony. 12 channel *Groups* were stacked to form a new baseband. These groups themselves were stacked with other groups to form *Supergroups*, and so on. Pilot tones were introduced to form the basis of automatic gain control for this complex system of up to 1800 voice channels. (2400 in Europe) Although at the time of writing (1998) many of these systems are being replaced with optical fiber links there are still many operating links throughout the world.

What Sparked My Insight?

Spark your educational insight @

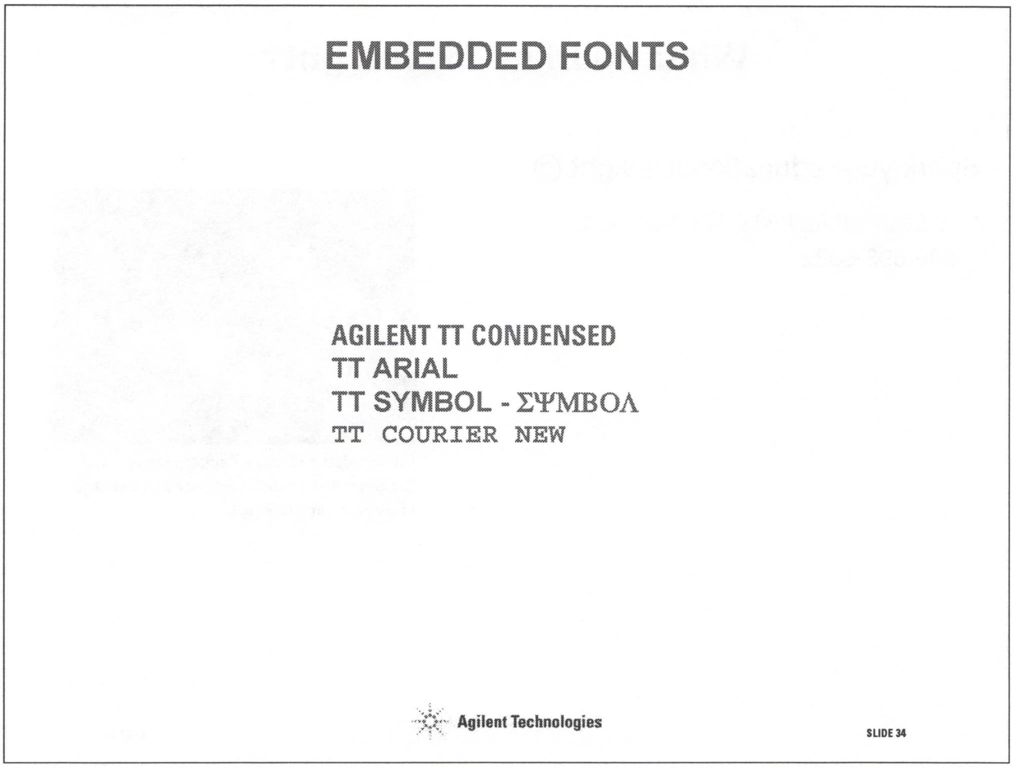
- www.agilent.com/find/tmeducation
- 800-593-6632



Bill Hewlett and Dave Packard founders of the Hewlett-Packard company and heritage of Agilent Technologies.

Agilent Technologies

SLIDE 33

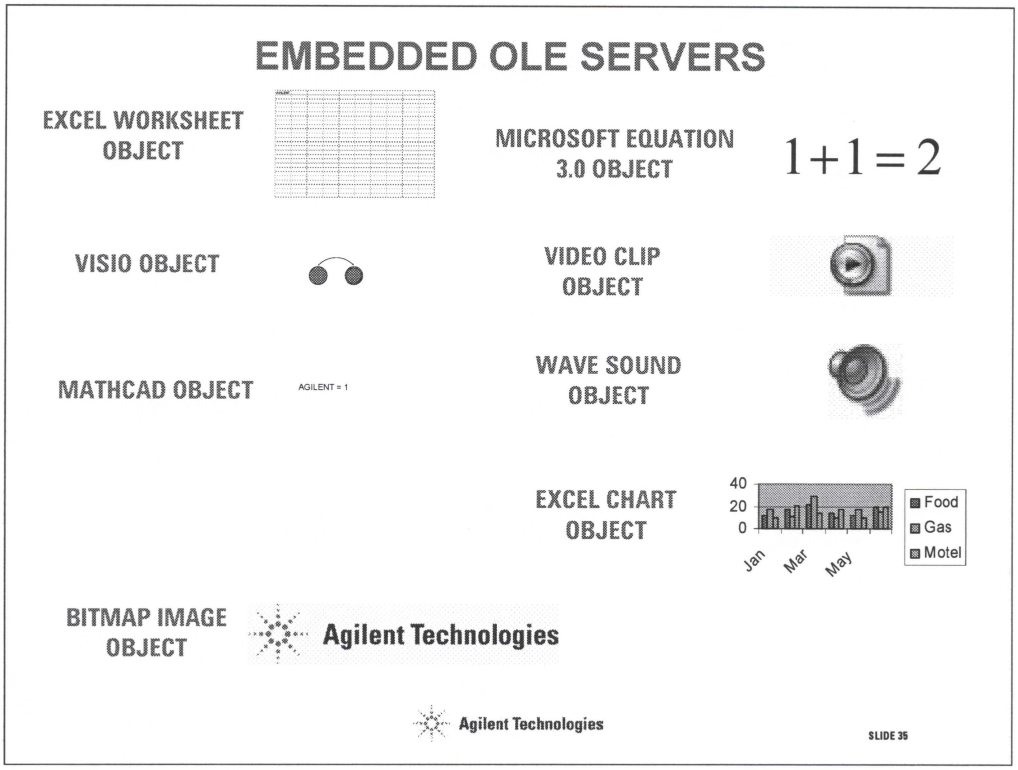


THIS SLIDE SHOWS THE EMBEDDED FONTS IN TEMPLATE.

TEMPLATE USERS ARE NOT ABSOLUTELY REQUIRED TO USE ONLY THE EMBEDDED FONTS, BUT IT IS RECOMMENDED.

IN PART, THE GENESIS OF THIS TEMPLATE, IS TO ELIMINATE THE FREQUENT PROBLEM OF SLIDES WITH FONTS THAT ARE NOT RESIDENT ON PC'S WITHIN THE USER COMMUNITY. BY LIMITING THE NUMBER OF FONTS AND EMBEDDING THEM INTO THE TEMPLATE, WE INTEND TO ELIMINATE THE PROBLEM OF SLIDE SETS THAT APPEAR CORRUPTED DUE TO LOCAL FONT SUBSTITUTION.

IF YOU REQUIRE AN ALTERNATIVE FONT FOR YOUR PRESENTATION, PLEASE EMBED THAT FONT IN YOUR PRESENTATION.



THIS SLIDE CONTAINS OLE OBJECTS AND CAUSES EMBEDDING OF THE RESPECTIVE OLE SERVERS IN THE TEMPLATE.

LEAVE THIS SLIDE AT THE VERY END OF THE PRESENTATION WHICH SHOULD FORCE THE OLE SERVERS TO REMAIN EMBEDDED IN YOUR PRESENTATION.

PLEASE EMBED ANY ADDITIONAL OLE SERVERS YOUR PRESENTATION MAY REQUIRE.

RF & Microwave Measurement Fundamentals

Frequency Modulation

Agilent Technologies

Objectives

Why FM?

How

Agilent Technologies

SLIDE 2

This section is about *angle modulation*. Angle modulation is the generic name for Frequency Modulation (FM) and Phase Modulation (PM). The non specialist is probably much more aware of FM because it is the alternative to AM on the family radio. Phase modulation has an important part to play in modern communications, especially in the field of digital communications.

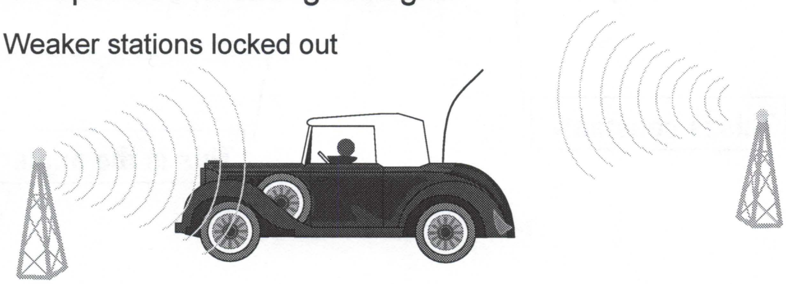
Why Bother with FM?

FM immune to amplitude variations

- Noise suppressed
- S/N improvement over AM

FM captures the strongest signal

- Weaker stations locked out



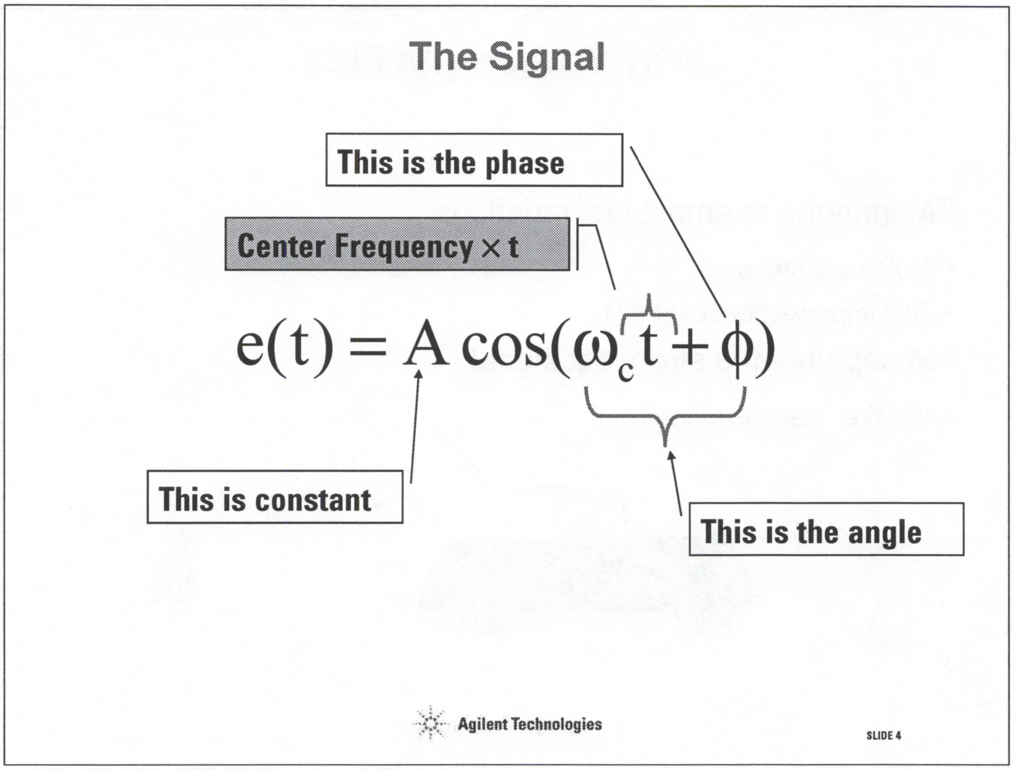
Agilent Technologies

SLIDE 3

Amplitude modulation can be very simple and for broadcast AM the receiving set can be very simple and cheaply replicated. The problem with AM is that interference called "static" would cause a severe degradation of the received signal specially when the signal was weak.

The inventor* of FM reasoned that if the intelligence of the signal, the baseband, modulated the frequency of the carrier, leaving the amplitude constant, then any interference would be ignored by the receive circuit. The first demonstrations of an FM broadcast were in 1933/1934, today we rely on FM for high fidelity listening.

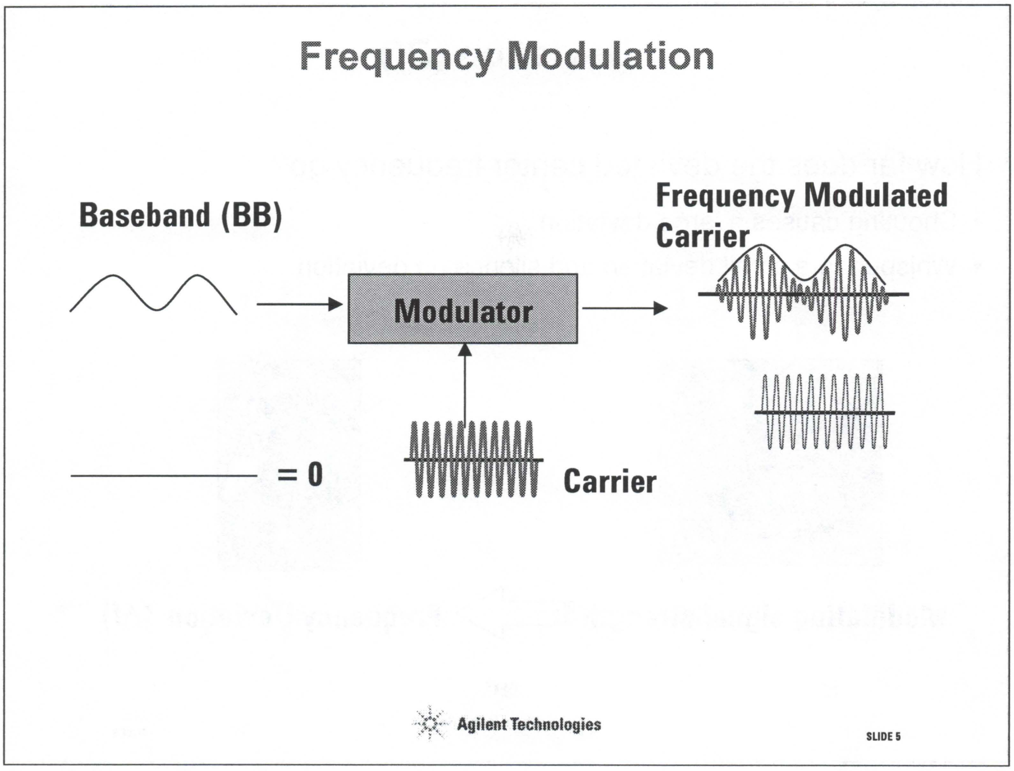
*Edwin Howard Armstrong. See "Empire of the Air - The Men who made Radio" Tom Lewis; Harper Collins.©1991



In angle modulation it is the angle($\omega_c t + \phi$), that is modulated. As in AM where the modulation signal varies a constant amplitude carrier, in Angle Modulation there is a constant *center frequency*. It can no longer be called a carrier because as we shall see under some conditions there is no signal there as in AM.

Frequency Modulation

In FM the phase (ϕ) becomes a function of time so that the center frequency is deviated by the modulating signal.



Frequency Modulation

This is the type of modulation used for the analog cellular standard and for commercial FM broadcast, and the audio part of broadcast and cable TV.

Frequency Modulation is not so intuitive as amplitude modulation. But quite a few facts my be deduced just by looking at the above drawing.

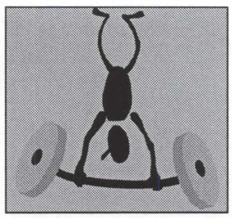
- The amplitude of the modulated wave does not change as the strength of the modulating signal changes.
- When the modulating signal is zero, the carrier is at the center frequency.

So what actually happens when the input modulating signal is made bigger?

Deviation (Df)

How far does the deviated center frequency go?

- Shouting causes a large deviation.
- Whispering a small deviation and silence no deviation.

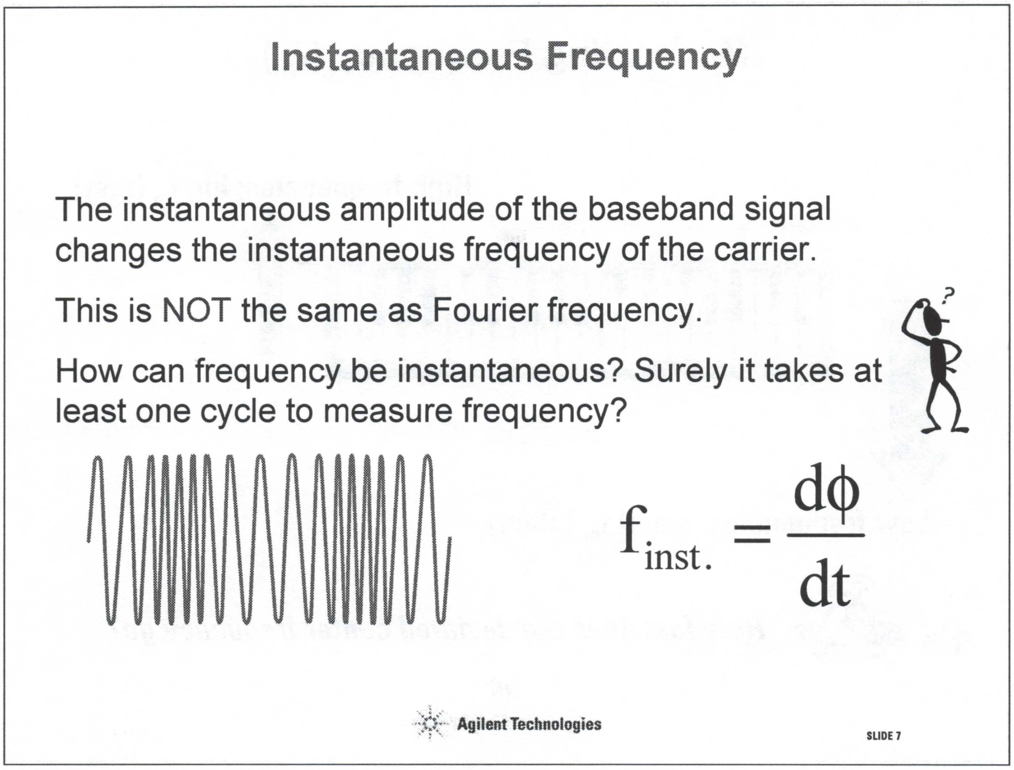


Modulating signal strength ⇨ Frequency Deviation (Δf)

Agilent Technologies

SLIDE 6

In frequency modulation the *strength* of the modulating signal *deviates* the center frequency from its central position. The deviation of an FM signal is expressed as the number of Hz. away from the signal's unmodulated central position. In all the analysis that follows the modulating signal will be a sinewave (tone modulation). The positive and negative peaks of a modulating sinewave will deviate the carrier either side of its central position by an amount called the peak deviation, Δf, this carrier is deviated $\pm \Delta f$ from the central position f_c.



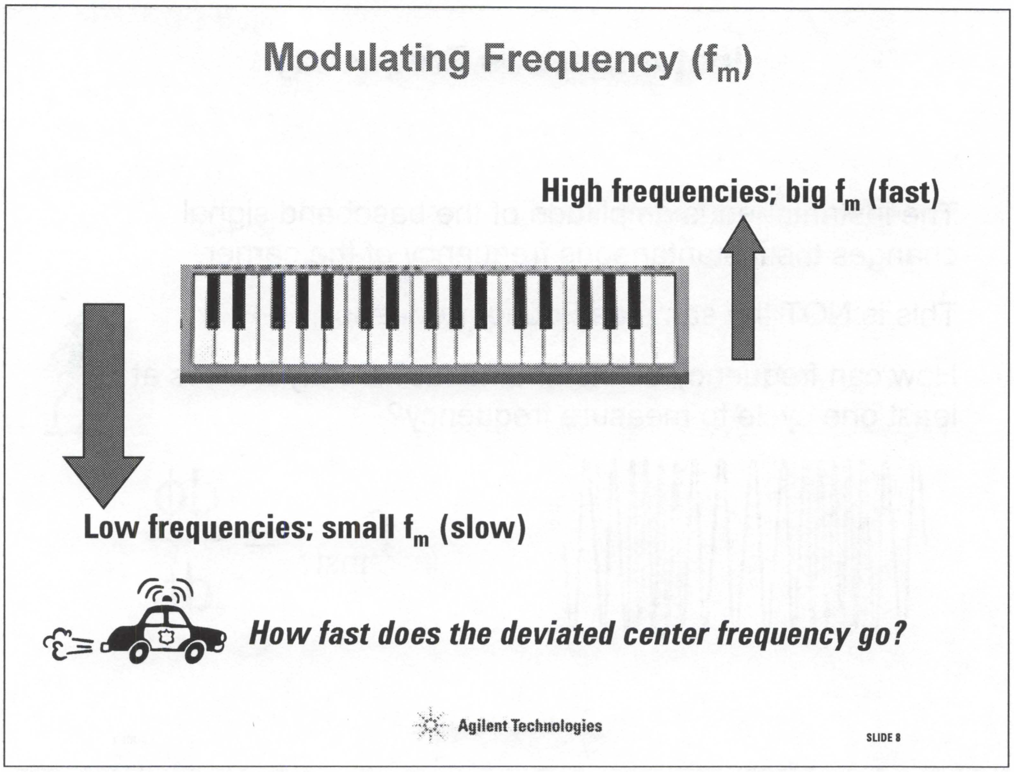
Although frequency is thought of in terms of cycles per second, the definition we shall take is, phase per unit time. Instead of doing numbers of complete 360 degree (2π radians) in a given time, the *rate of change of phase* will be our new definition.The units are still deg/sec or rads/sec. It is the speed of rotation of the phasor.

For example; if you are cycling at 15 mph your speed does not depend on cycling for 1 hour or going 15 miles. In physics we define velocity (v) as the rate of change of displacement (s).

$$v = \frac{ds}{dt}$$

So in signal theory:

$$f_{inst.} = \frac{d\phi}{dt}$$

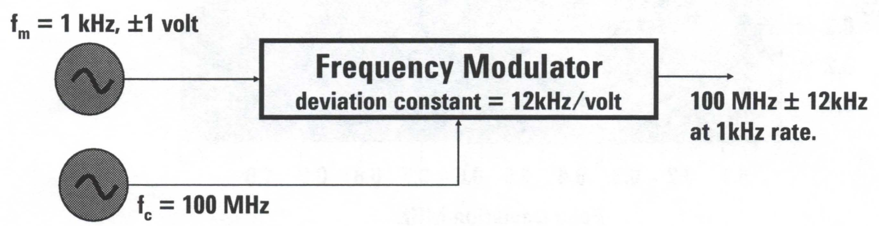


The modulating frequency (f_m) is proportional to the *tone* of the baseband signal. In FM both the strength and the frequency of the baseband signal is carried in the frequency/phase of the deviated carrier. Both f_m and Δf are measured in Hz. This is sometimes a source of confusion.

The Frequency Modulator

In Frequency Modulation, an RF signal is deviated in sympathy with the modulating signal.

- A 1kHz, 1volt peak tone, frequency modulating a 100MHz carrier will deviate it at a 1kHz rate and with a peak deviation proportional to 1volt.



Agilent Technologies

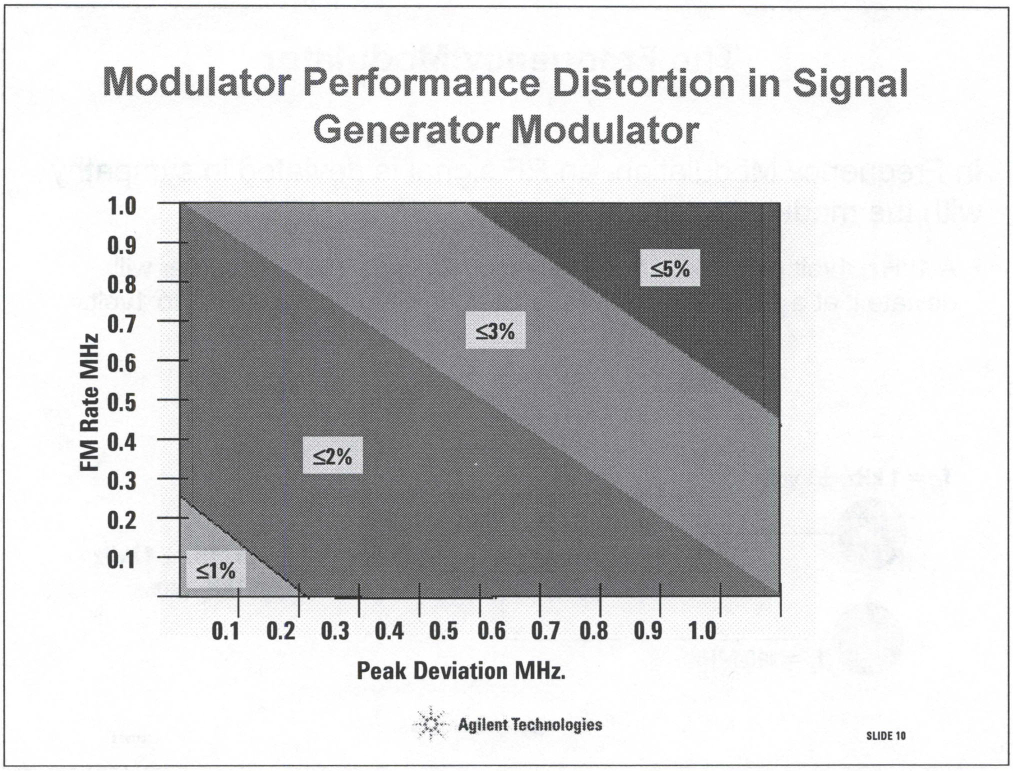
SLIDE 9

It is the *magnitude* of the modulating signal that *deviates* the carrier away from its center frequency. This example shows this relationship between the magnitude and frequency of the modulating signal and the resulting modulated signal.

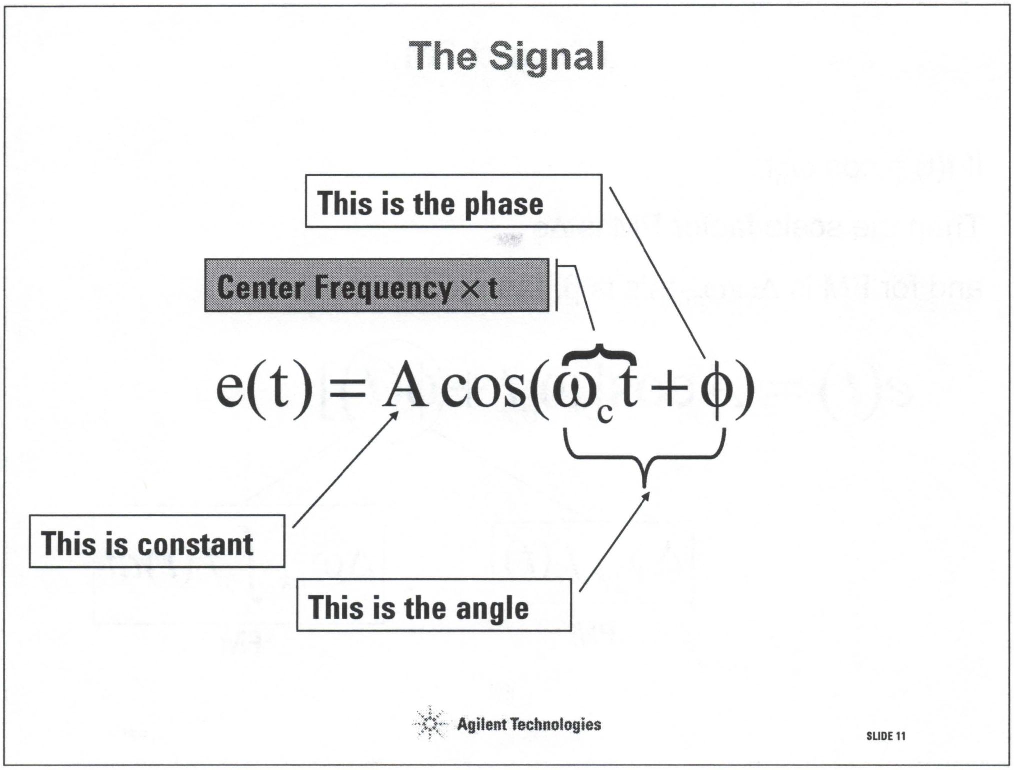
A key specification in FM is the deviation constant of the modulator which quantifies how the strength of the input signal becomes deviation in frequency.

Another key specification of the modulator would be its bandwidth. A frequency modulator used for audio may have a bandwidth from 100Hz to 15kHz. Such a band is referred to as AC coupled, if the modulator were DC coupled its lower frequency would be 0 Hz or DC.

A modulator can generate a maximum deviation for a given modulation frequency, for a given distortion.



A good signal generator will be specified to accept a range of modulating frequencies at given deviations.



In angle modulation it is the angle($\omega_c t + \phi$), that is modulated. As in AM where the modulation signal varies a constant amplitude carrier, in angle modulation the amplitude is constant and the center frequency angle is modulated. The angle modulated signal has no constant carrier, the signal that is present when there is no modulation, will decrease in amplitude as the modulation is increased and can disappear under some conditions.

PM and FM

If $f(t) = \cos \omega_m t$

Then the scale factor PM is $\Delta\phi$

and for FM is $\Delta\omega/\omega_m$ this is β, the modulation index.

$$e(t) = A\cos[\omega_c t + \phi(t)]$$

$$\boxed{\Delta\phi_{pk} f(t)}$$

PM

$$\boxed{\Delta\omega_{pk} \int f(t)dt}$$

FM

Agilent Technologies

SLIDE 12

The variable part of the expression for angle modulation is in the "argument" of the sin (or cosine) function. This part is expressed differently for phase and frequency modulation. In phase modulation the phase is proportional to the instantaneous strength of the modulating signal, in frequency modulation, since frequency is the time rate of change of phase, the phase is the integral of the time function.

In FM the ratio $\Delta\omega/\omega_m$ the deviation ratio is also known as the modulation index. An analogy to AM.

Using a Tone Modulation

$$e(t) = A\cos[\omega_c t + \phi(t)]$$

If $\phi(t) = \beta \cos\omega_m t$

We get cosine of a cosine!

$$e(t) = A\cos[\omega_c t + \beta\cos\omega_m t]$$

Agilent Technologies

SLIDE 13

In going further with this, using a sinusoidal time function as the modulating signal, the expression becomes complicated by having a cosine of a cosine. There is a solution to this which involves a new function, the **Bessel** function. Bessel functions are a family of functions which when graphed look like cosine or sines with a decaying envelope.

Bessel functions are denoted by J with a subscript and an argument, for example:

$J_0(\beta)$ is the zero order Bessel function at β radians

Lots of Sidebands!

$$\boxed{\begin{aligned}&\cos(\omega t + \beta \cos \omega_m t) = \\ &\cos \omega t \cos(\beta \cos \omega_m t) - \sin \omega t \sin(\beta \cos \omega_m t)\end{aligned}}$$

- **These products yield series involving Bessel functions:**

Even terms: $\cos(\beta \cos x) = J_0(\beta) + 2\sum_{n=1}^{\infty} (-1)^n J_{2n}(\beta) \cos 2nx$

Odd terms: $\sin(\beta \sin x) = 2\sum_{n=1}^{\infty} J_{2n-1}(\beta) \sin(2n-1)x$

Agilent Technologies

SLIDE 14

This simple tone modulation, produces a series of discrete tones all separated by the modulating tone frequency. It is not surprising to reflect that when FM was first suggested this infinite number of spectral components was seen as not being practical.

FM Spectral Components

$$e_{FM} = J_0(\beta)\cos\omega_c t$$

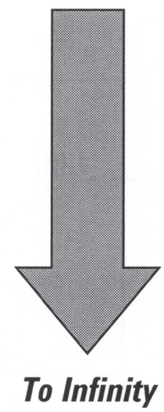
$$+J_1(\beta)[\cos(\omega_c+\omega_m)t-\cos(\omega_c-\omega_m)t]$$

$$+J_2(\beta)[\cos(\omega_c+2\omega_m)t-\cos(\omega_c-2\omega_m)t]$$

$$+J_3(\beta)[\cos(\omega_c+3\omega_m)t-\cos(\omega_c-3\omega_m)t]$$

$$+J_4(\beta)[\cos(\omega_c+4\omega_m)t-\cos(\omega_c-4\omega_m)t]$$

$$+\ldots$$

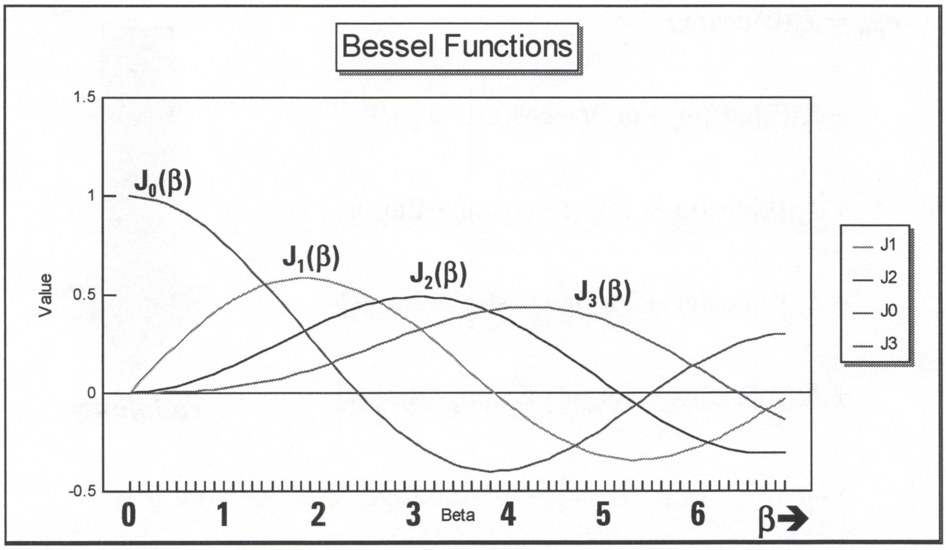


Agilent Technologies

SLIDE 15

The carrier is multiplied by $J_0(\beta)$, the other terms are sidetones, The first pair of sidetones, multiplied by $J_1(\beta)$ are at the sum and difference frequencies, somewhat like AM, but there are many more.

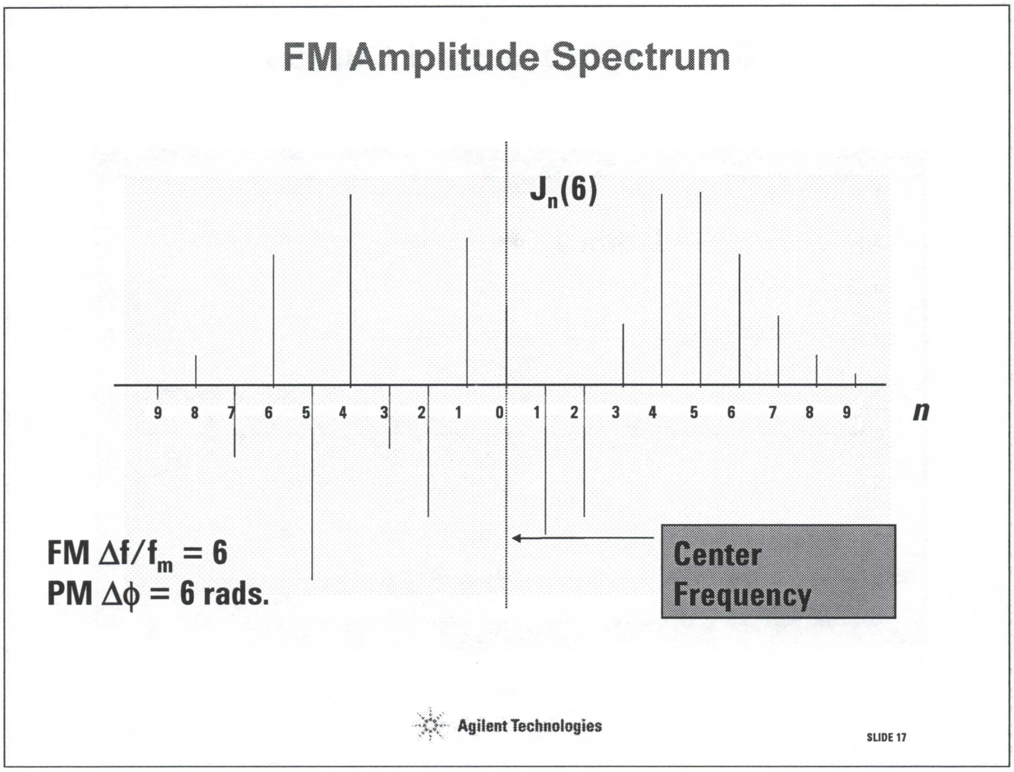
What is J? What are Bessel Functions?



Agilent Technologies

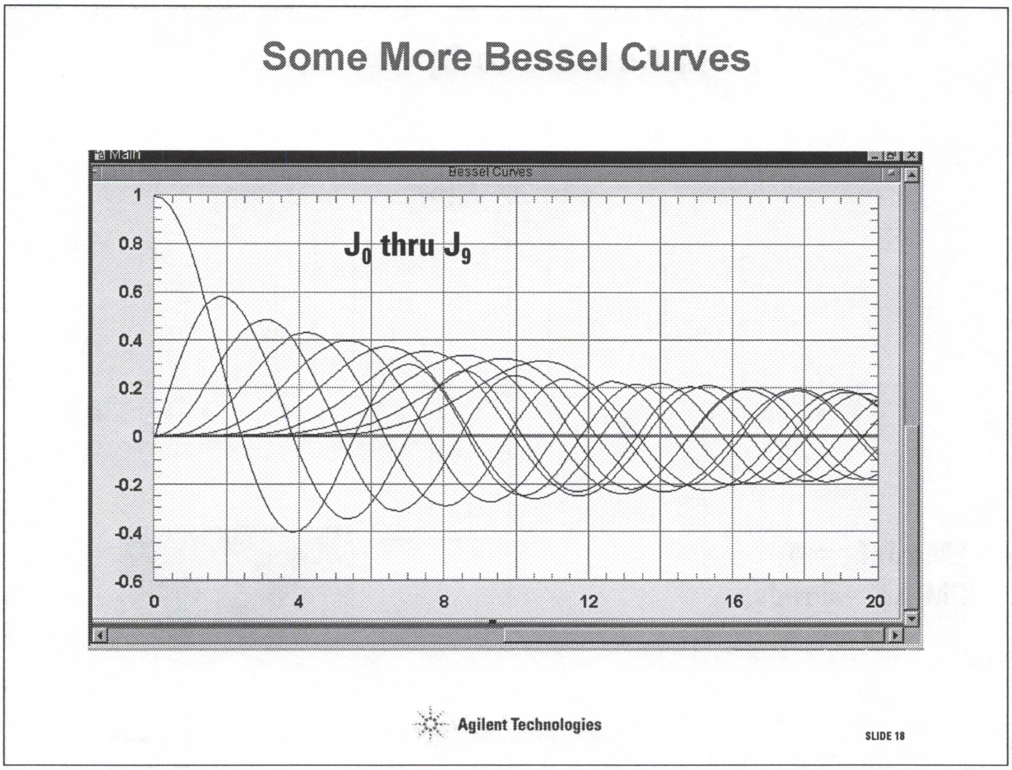
SLIDE 16

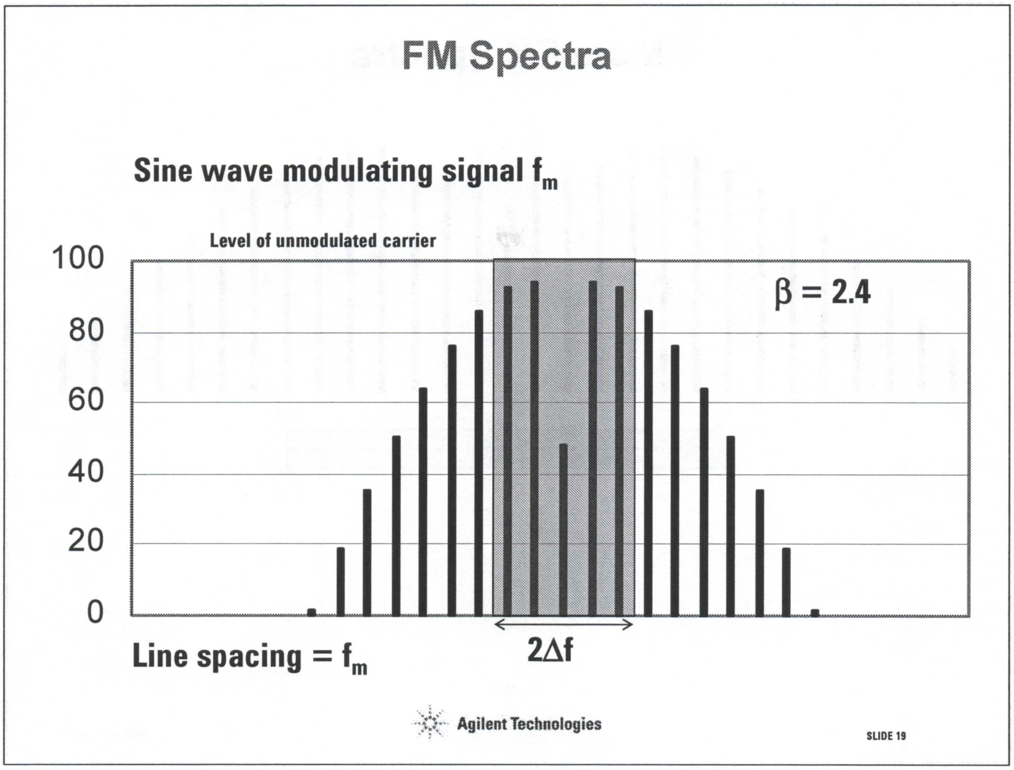
The Bessel functions look like damped sinewaves. All these and many more, are the coefficients of the sideband pairs produced in FM.



This is the spectrum of an FM signal showing most of the significant J components. On a spectrum analyzer the components would be all positive as only the absolute values would be displayed. Most important note that the components get smaller as *n* gets larger. Even though there are an infinite number of components, only a finite number are large enough to be significant.

The sign of the function is the relative phase at time t=0. Imagine them all rotating in and out of the plane of the paper at speeds proportional to their harmonic numbers.


Some More Bessel Curves
Main
Bessel Curves
J_0 thru J_9
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
0
4
8
12
16
20
Agilent Technologies
SLIDE 18



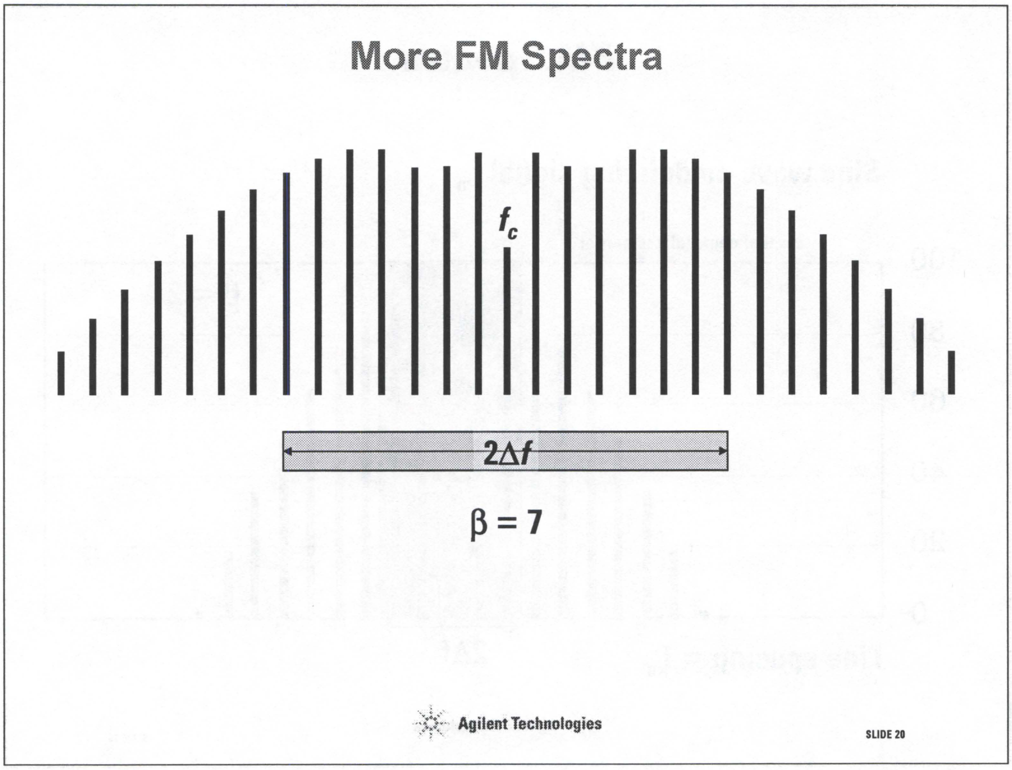
From our first slide on FM we decided that the power of the transmitted signal never changes. If the spectrum of an FM signal is displayed all the lines are *below* the level of the unmodulated carrier, the scale in our illustration is a dB scale. The sum of the powers in each spectral line would equal the power of the unmodulated carrier. Notice how the center component which is at the frequency of the unmodulated carrier is noticeably low.

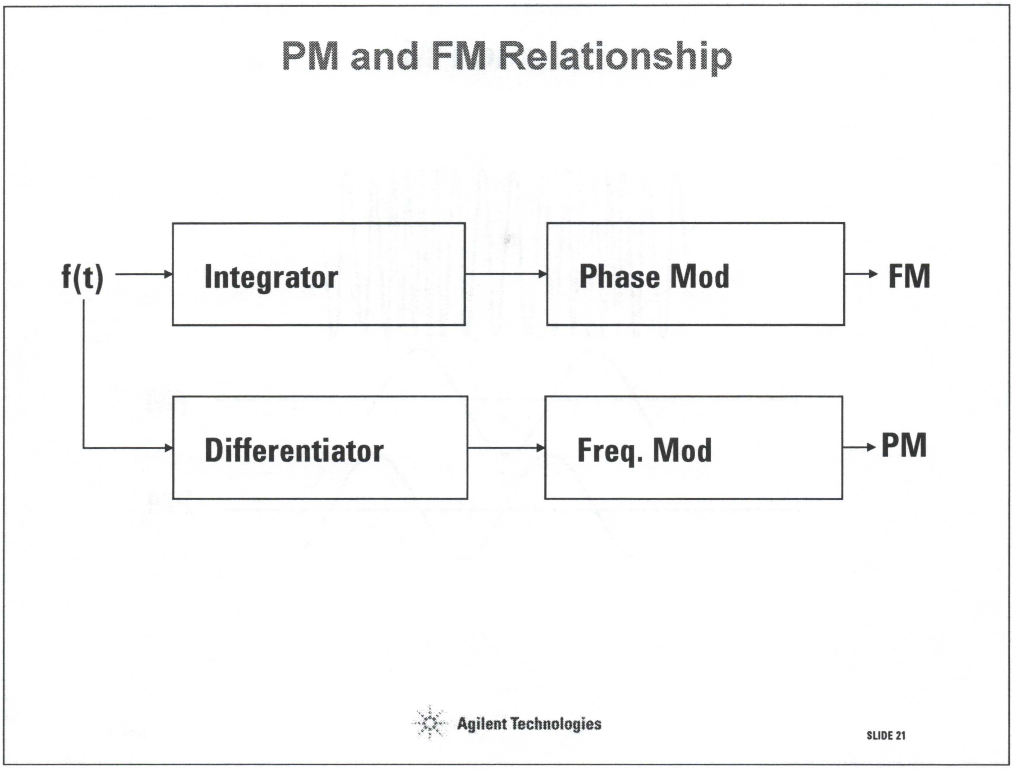
Note: ***the spectrum extends well beyond the bounds of the deviation.***

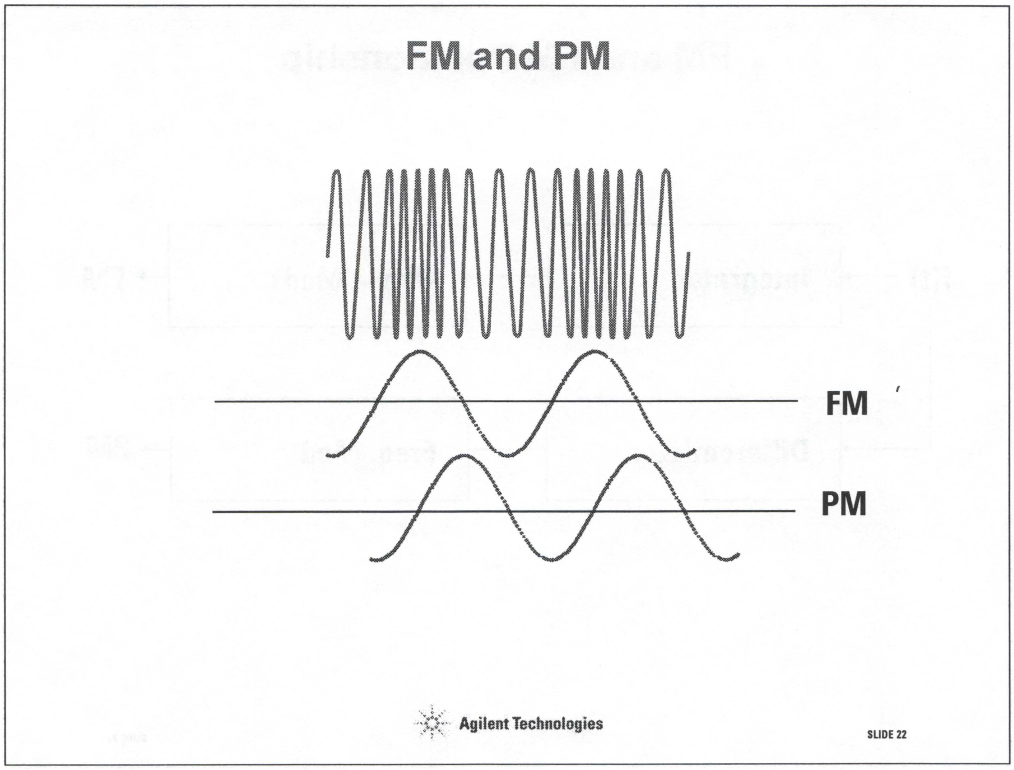
For example in an analog cellular (AMPS) system the voice channel peak deviation is 12kHz, the bandwidth needed for the channel is 30kHz.

$$\beta = \frac{\Delta f}{f_m}$$

$\beta = 2.4$
Carrier goes to ∅.


More FM Spectra
f_c
$2\Delta f$
$\beta = 7$
Agilent Technologies
SLIDE 20


PM and FM Relationship
f(t)
Integrator
Phase Mod
FM
Differentiator
Freq. Mod
PM
Agilent Technologies
SLIDE 21



With tone modulation the resulting wave from the FM process is the same as a phase shifted PM process. The spectrum of PM is not distinguishable from FM.

Some Spectra - Single Tone Modulation

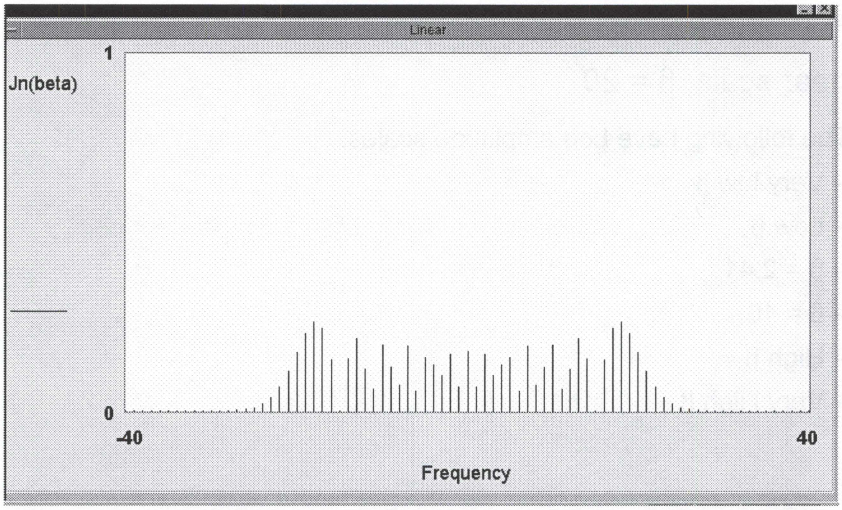
Linear scale, $\beta = 20$

- The following have Log amplitude scales:
  - Very low β
  - Low β
  - $\beta = 2.41$
  - $\beta = 10$
  - High β
  - Very High β

Agilent Technologies

SLIDE 23

Linear Scale b = 20



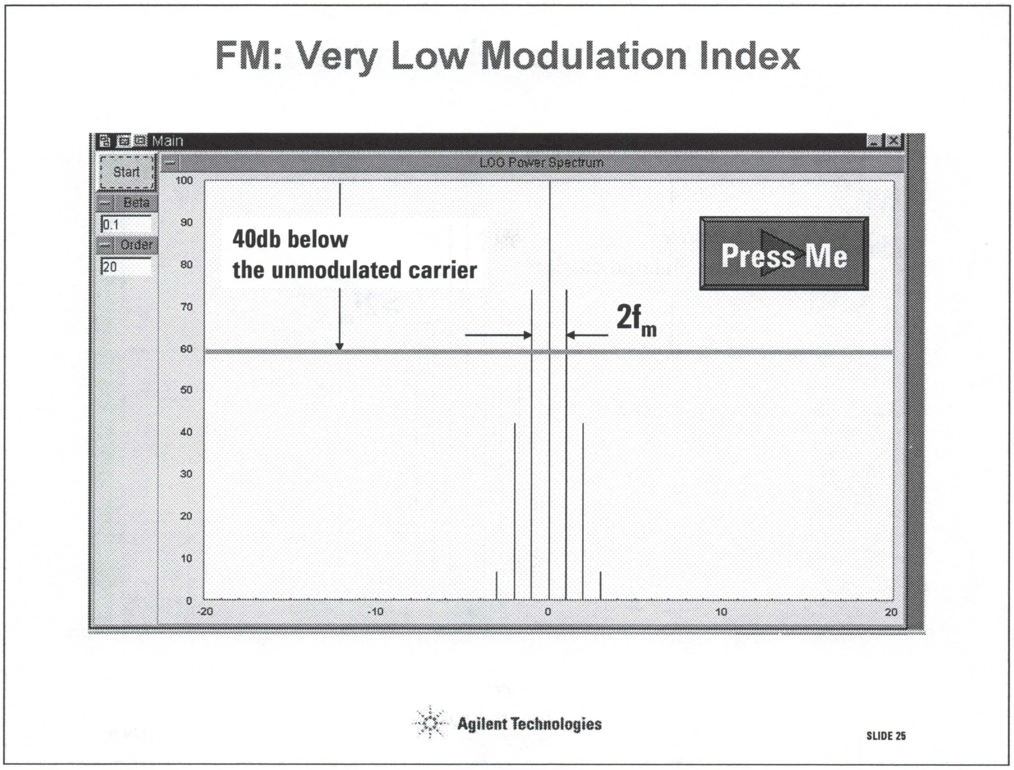
Agilent Technologies

SLIDE 24

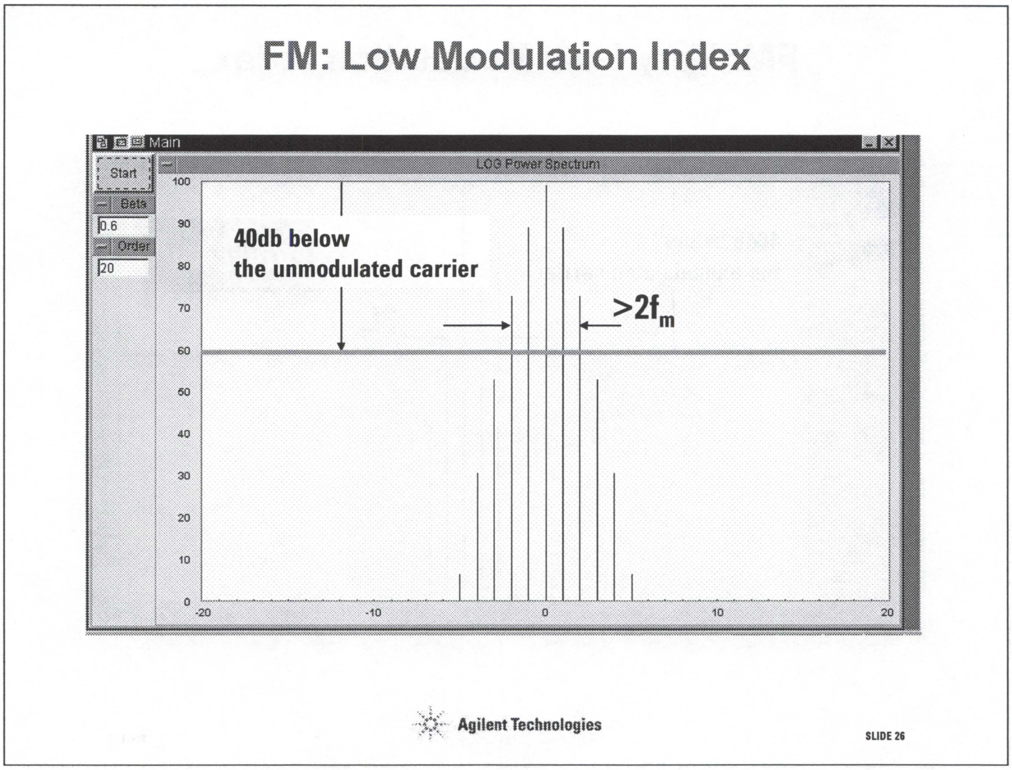
With the linear scale it is easier to visualize that the sum of the powers of the spectral lines is the power of the unmodulated carrier.

In these examples ***V*** (unmodulated) = 1

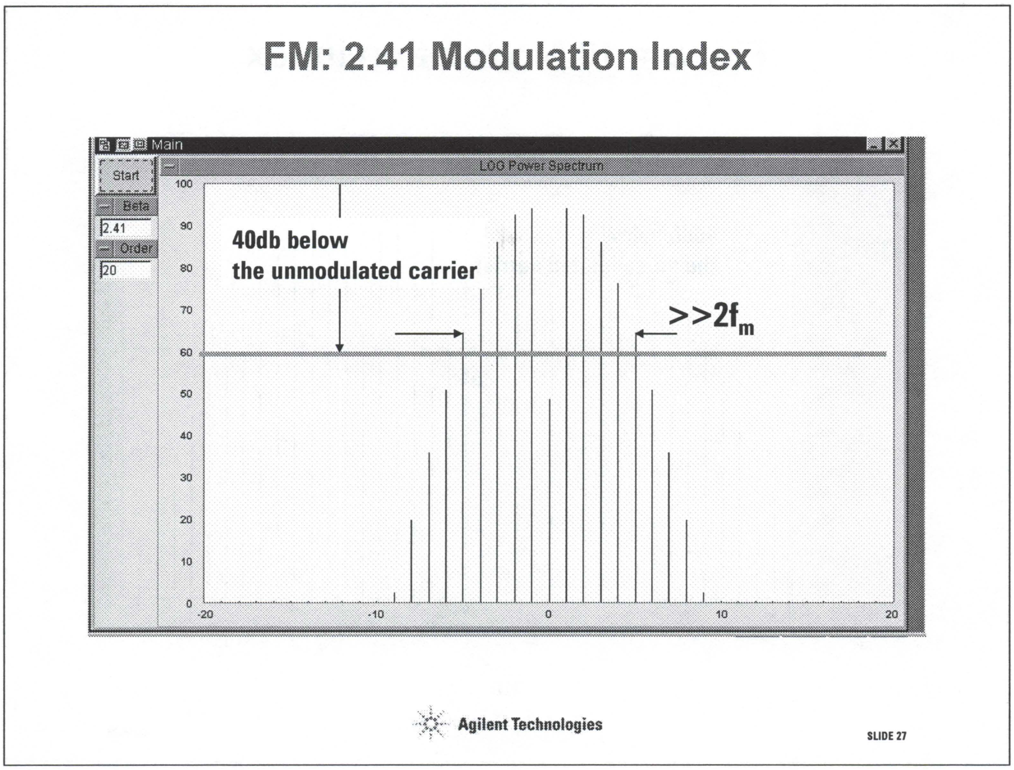
$$\sum_{n=-\infty}^{n=+\infty} \left(V_n\right)^2 = 1$$



With a very low modulation index there are just two significant sidebands. This almost looks like an AM spectrum. The occupied bandwidth is twice the modulating frequency, f_m.

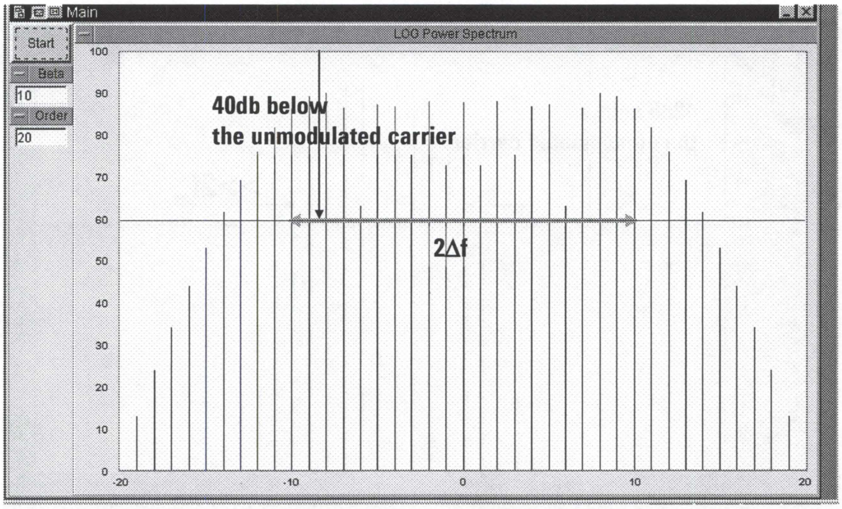


As deviation increases the number of significant sidebands increases. Here β is 0.6



The deviation has been increased further so that β = 2.41. Notice that the signal at the center has dropped significantly. This is very close to the First Bessel Null condition β = 2.4048 ...

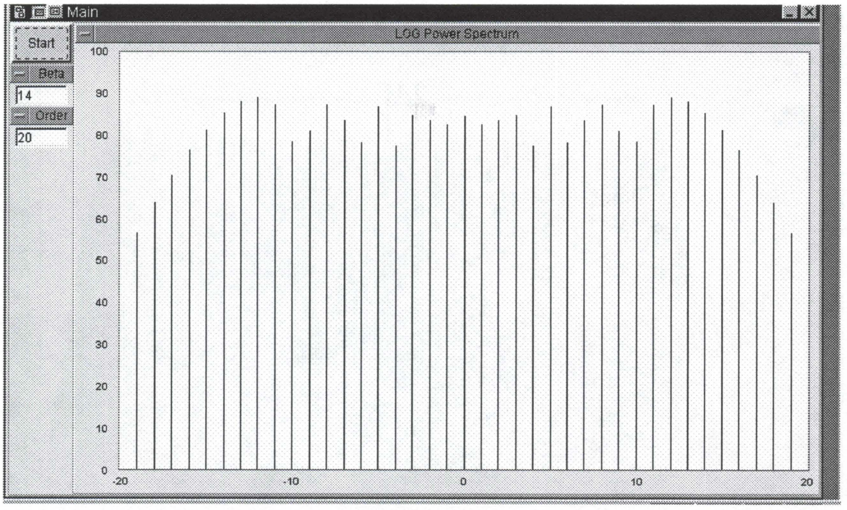
FM: b = 10 Modulation Index



Agilent Technologies

SLIDE 28

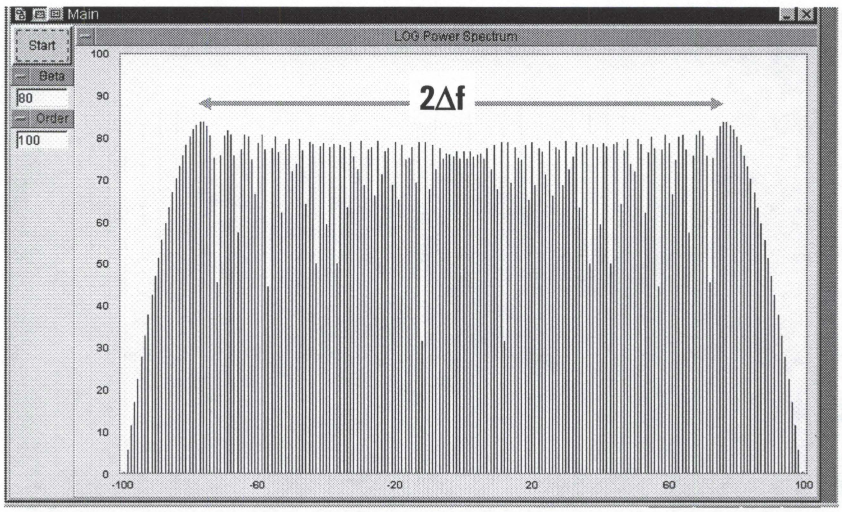
FM: Higher Modulation Index



Agilent Technologies

SLIDE 29

FM: Very High Modulation Index



Agilent Technologies

SLIDE 30

When the modulation has a very large index, here β is 80, the occupied spectrum is almost $2\Delta f$.

Required Bandwidth

Index	N	BW in f_m	BW in Δf
0.1	2	2	20
0.3	4	4	14
0.5	4	4	8
1.0	6	6	6
2.0	8	8	4
5.0	16	16	3.2
10.0	28	28	2.8
30.0	70	70	2.3

BW = $2f_m$

BW → $2\Delta f$

Agilent Technologies

SLIDE 31

This chart shows the BW (occupied spectrum) stated in multiples of modulating frequency, f_m, and multiples of the deviation, Δf, for increasing deviation ratio $\Delta f/f_m$. At small deviation ratios the BW is about $2\times f_m$, and at very large deviations this approaches $2\times\ \Delta f$.

Carson's Rule for FM

$$\text{Occupied Spectrum} \approx 2(f_m + \Delta f)$$

Agilent Technologies

SLIDE 32

Unlike AM the estimation of bandwidth in an FM situation is not exactly calculable. There are various estimates of transmission bandwidth, the most widely used is Carson's rule stated above. If, for example, Carson's rule is applied to US broadcast FM the rule underestimates the bandwith allowed.

Δf (max) = 75kHz, f_m (max) = 15kHz : B = 2(75 +15)kHz = 180kHz the bandwidth of an FM radio is usually 200kHz.

FM Use

Radio: 87.9 to 107.9 MHz 200 kHz channels

- 15kHz audio BW max deviation 75kHz

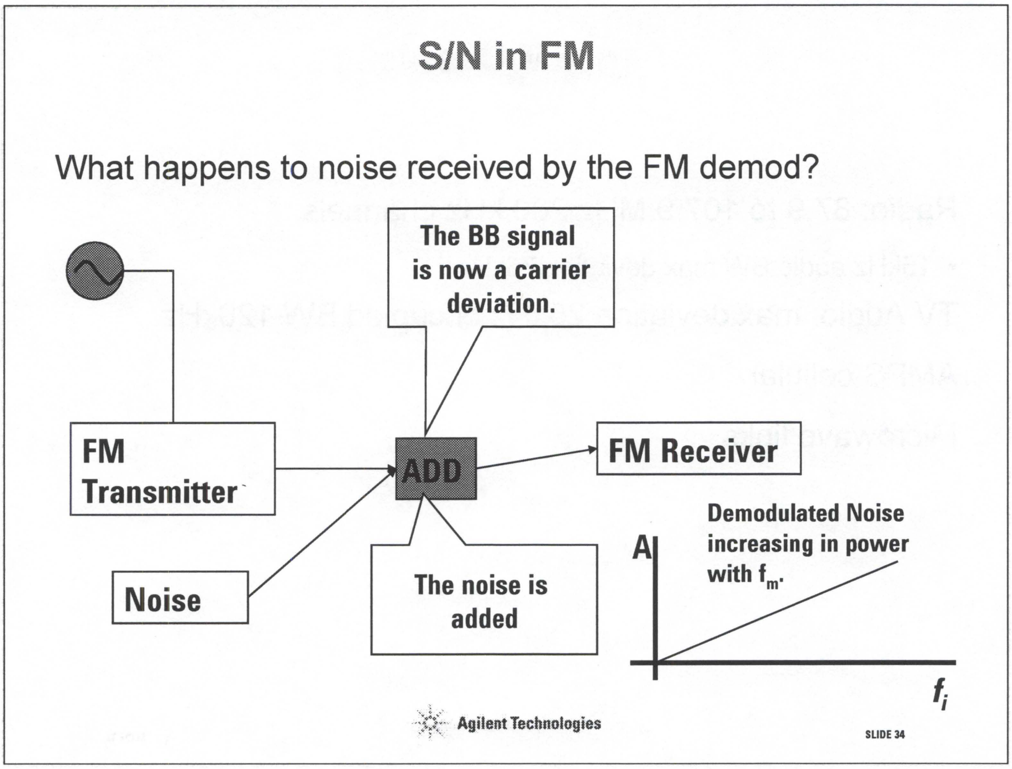
TV Audio: max deviation 25kHz occupied BW 120kHz.

AMPS cellular

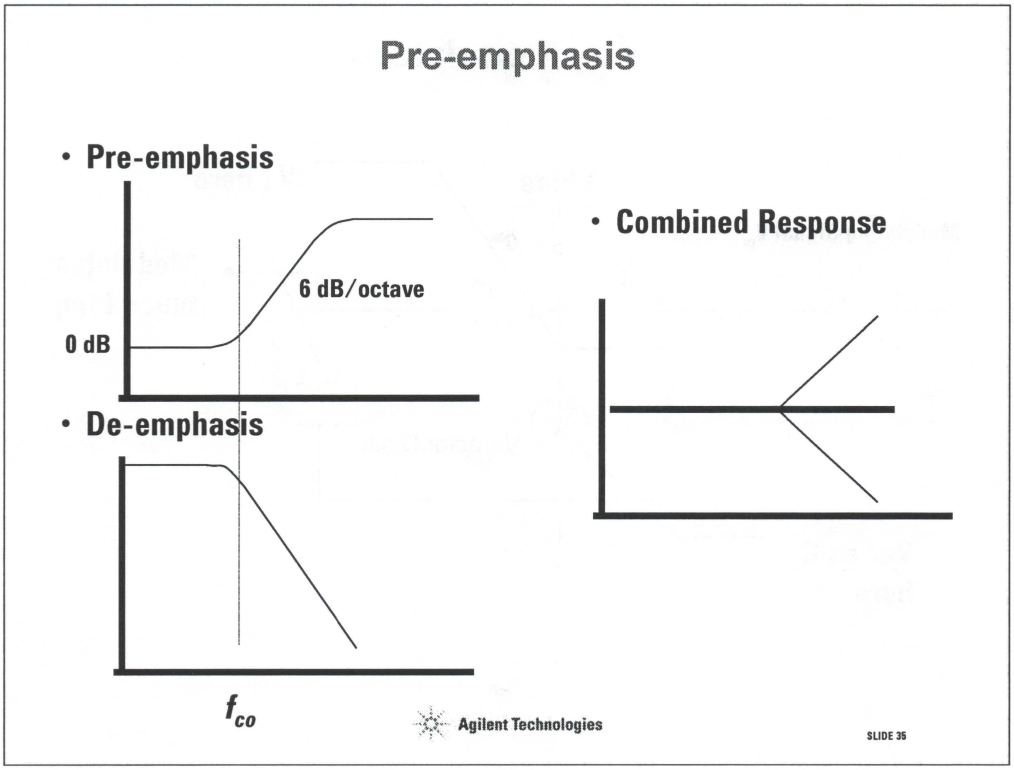
Microwave links

Agilent Technologies

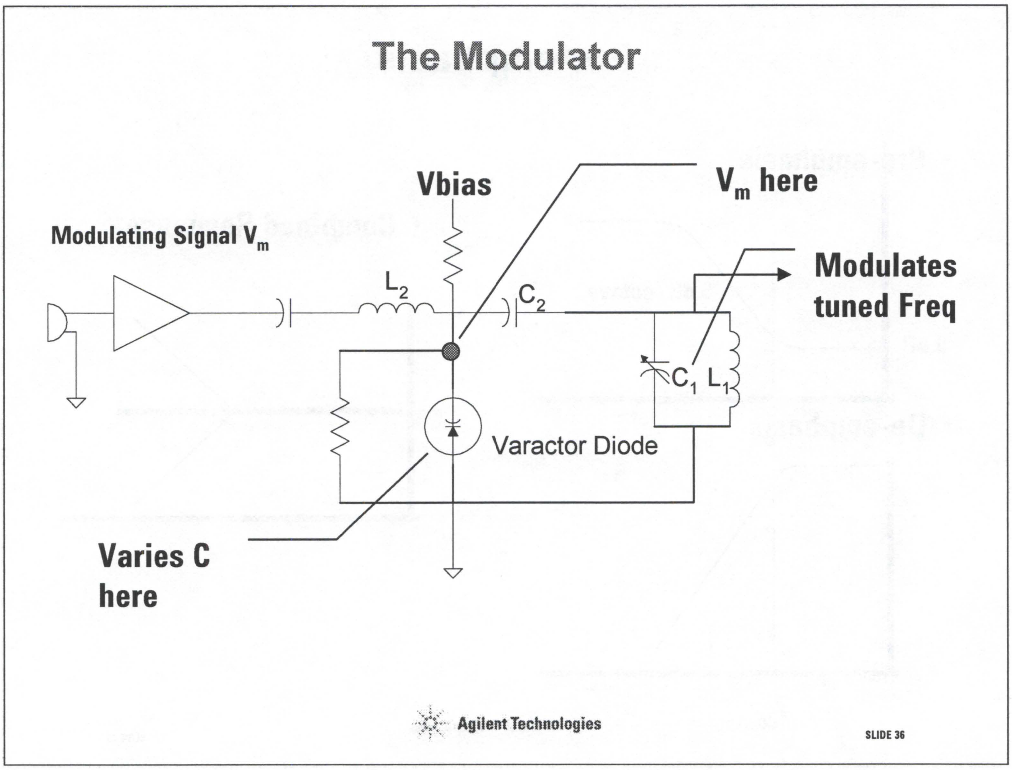
SLIDE 33



Interference picked up by the FM receiver which was not part of the required signal will have an effect when demodulated which increases with the frequency of the interference. This is why in FM broadcast systems the baseband signal is pre-emphasized at 6dB/octave from about 2kHz. This ensures a good S/N ratio of the demodulated signal. The pre-distorted signal must be then de-emphasized.

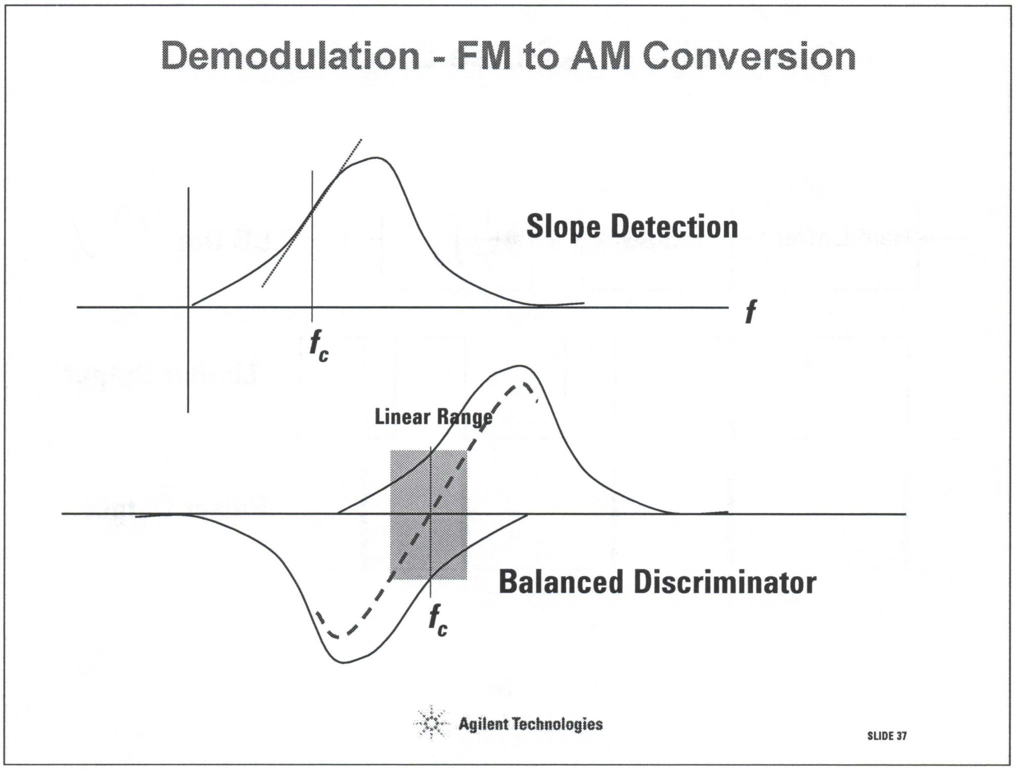


This shows the combined effect of re-emphasis and de-emphasis on the required signal.

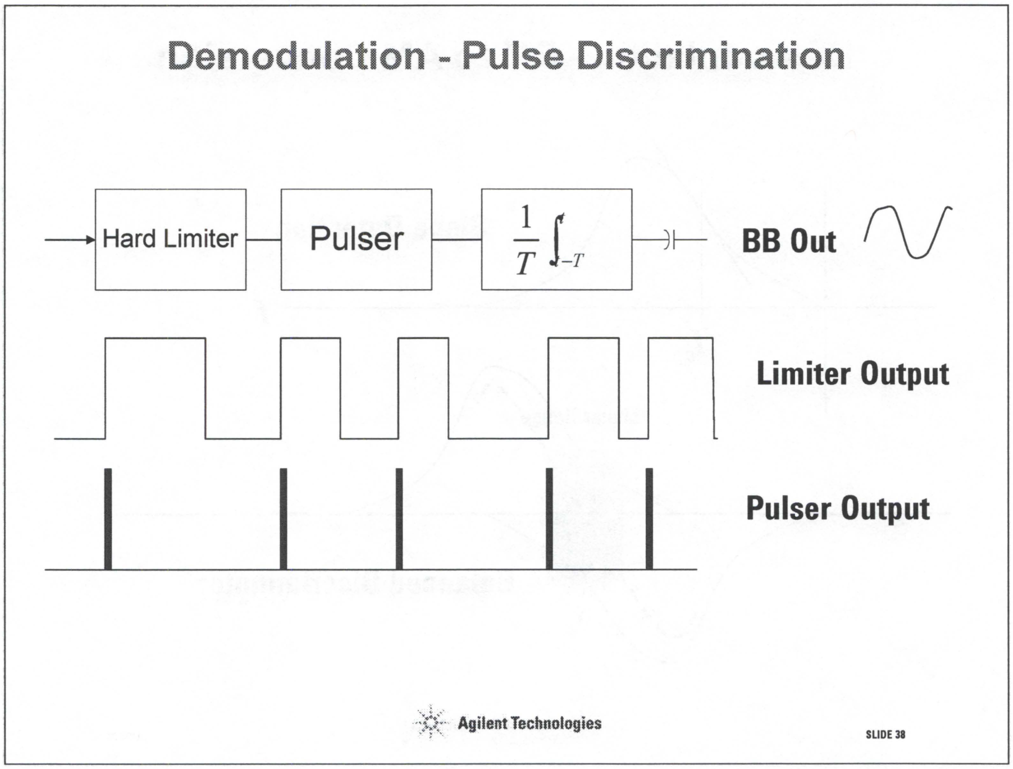


There are different ways of causing a frequency modulated signal, here is shown just one of them. The key device is the Varactor diode, which is a special junction diode which is optimized to change capacitance as the bias voltage changes.

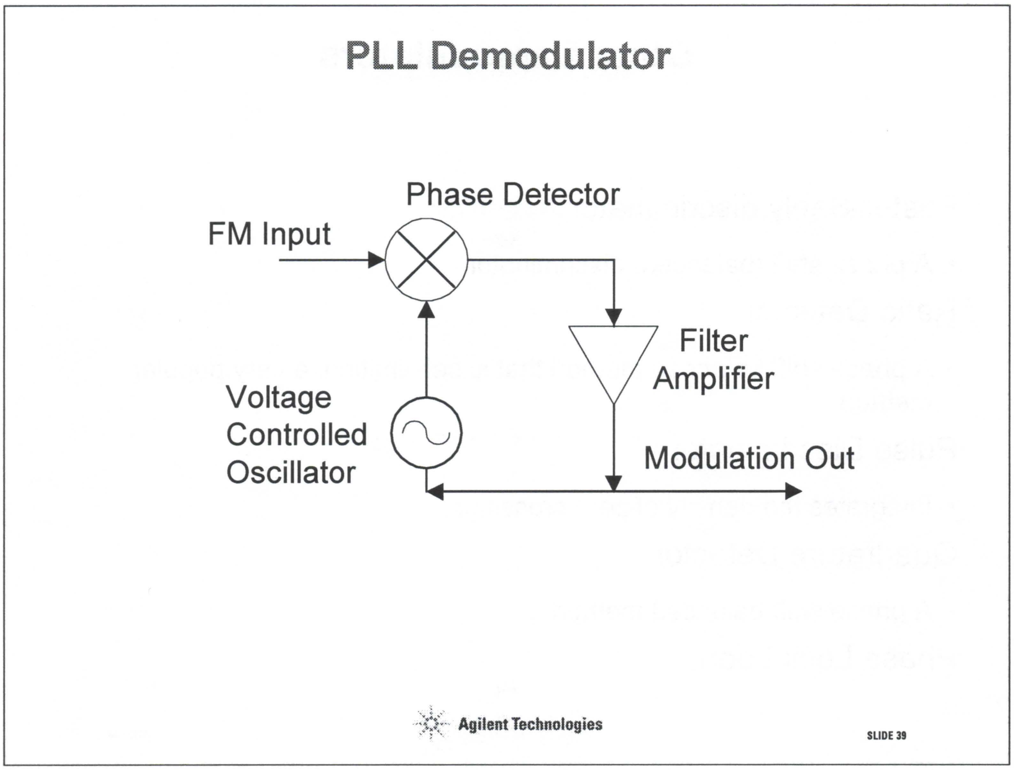
L_2 has a high reactance and C_2 a low reactance to the RF oscillations produced by the resonant circuit C_1,L_1. The diode capacitance is in parallel with the tuned circuit, and varies its tuned frequency in sympathy with the modulating voltage Vm.



The demodulation of FM may be done with a discriminator which essentially converts FM deviation into AM. The linear operation of the above would be restricted to a small range close to f_c.



In the Pulse Count Discriminator, the zero crossings are used to create a pulse position signal, the density of the pulses represent the baseband signal.



The Phase Lock Loop Demodulator

With no modulation the VCO is at the center frequency, the phase detector output will be zero.

As the input deviates the an error voltage is produced which pushes the VCO frequency to match this new input.

This VCO is a frequency modulator which is tracking the input modulation, the error voltage is a replica of the original modulating signal.

Other Demodulators

Foster-Seely discriminator

- A phase shift (balanced) discriminator

Ratio Detector

- A phase shift balanced method that is self limiting, a very popular method.

Pulse Discriminator

- Integrates the density of zero crossings.

Quadrature Detector

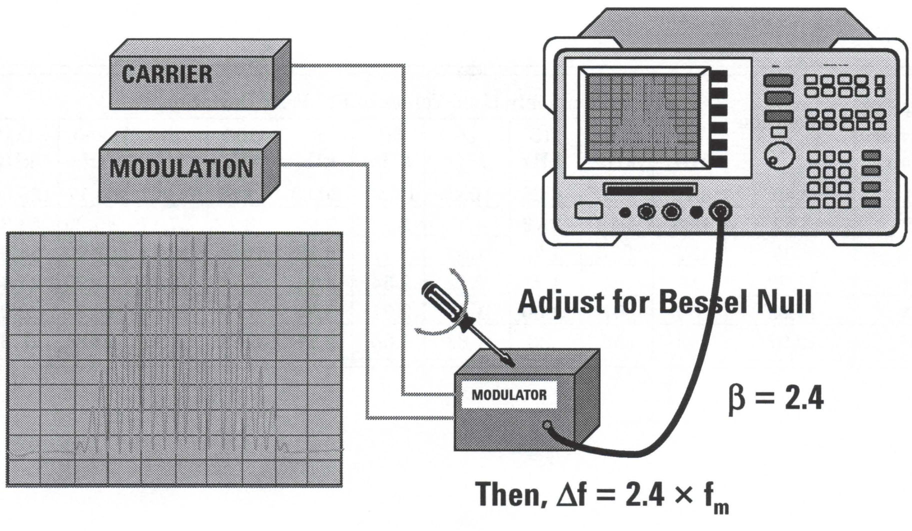
- A phase shift balanced method.

Phase Lock Loop

Agilent Technologies

SLIDE 40

Calibrating Deviation with a Spectrum Analyzer



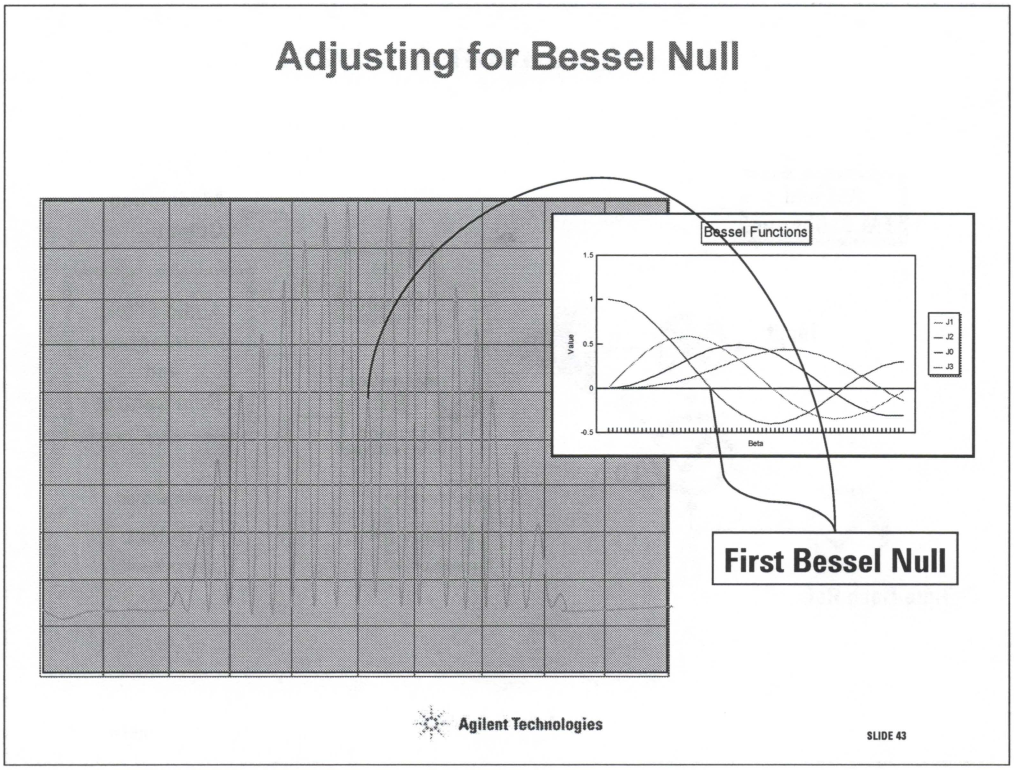
If the modulation is a single tone the measurement of the modulating frequency f_m on a spectrum analyzer is a matter of measuring the line spacing, or it could be measured with a frequency counter. But the deviation is not a simple measurement, but using our knowledge of the Bessel functions, and knowing that $J_o(\beta)$ will be zero at β = 2.4048. We can adjust the modulator and watch for the disappearance of the center signal.

Various Carrier Nulls

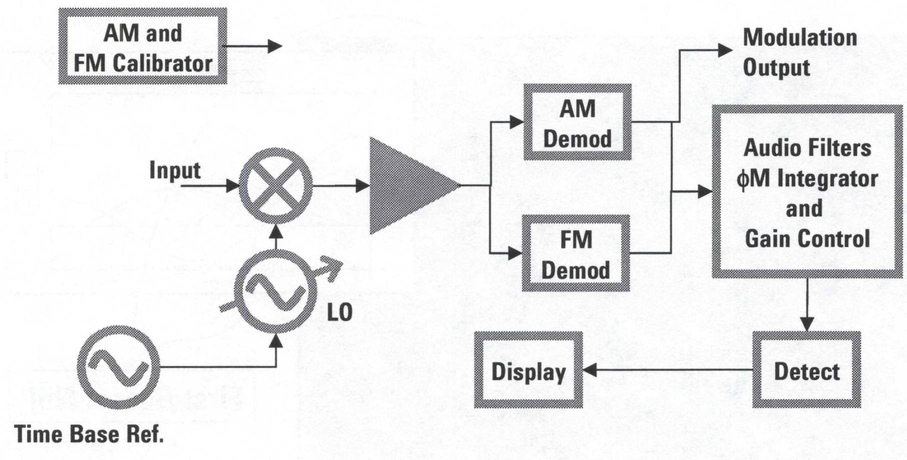
Order of Carrier zero	Modulation Index	7.5 kHz	Commonly Used Values of FM Peak Deviation 10 kHz	15 kHz	25 kHz	30 kHz	50 kHz	100 kHz	150 kHz	250 kHz	300 kHz
1	2.40	3.12	4.16	6.25	10.42	12.50	20.83	41.67	62.501	04.17	125.00
2	5.52	1.36	1.18	2.72	4.53	5.43	9.06	18.12	27.17	45.29	54.35
3	8.65	.87	1.16	1.73	2.89	3.47	5.78	11.56	17.34	28.90	34.68
4	11.79	.66	.85	1.27	2.12	2.54	4.24	8.48	12.72	21.20	25.45
5	14.93	.50	.67	1.00	1.67	2.01	3.35	6.70	10.05	16.74	20.09
6	18.07	.42	.55	.83	1.88	1.66	2.77	5.53	8.30	13.84	16.60

Agilent Technologies

SLIDE 42


Adjusting for Bessel Null
Bessel Functions
1.5
1
0.5
0
-0.5
Value
Beta
J1
J2
J0
J3
First Bessel Null
Agilent Technologies
SLIDE 43

The Modulation Analyzer



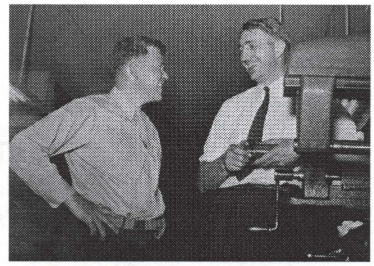
Special equipment to measure modulation in general and FM in particular depends on the classic heterodyne down conversion followed by a reference FM demodulator. In the HP8901B modulation analyzer the FM demod is called a charge count discriminator which has been designed to have both low noise and highly linear performance.

Phase demodulation is achieved by integrating the output of the FM discriminator.

What Sparked My Insight?

Spark your educational insight @

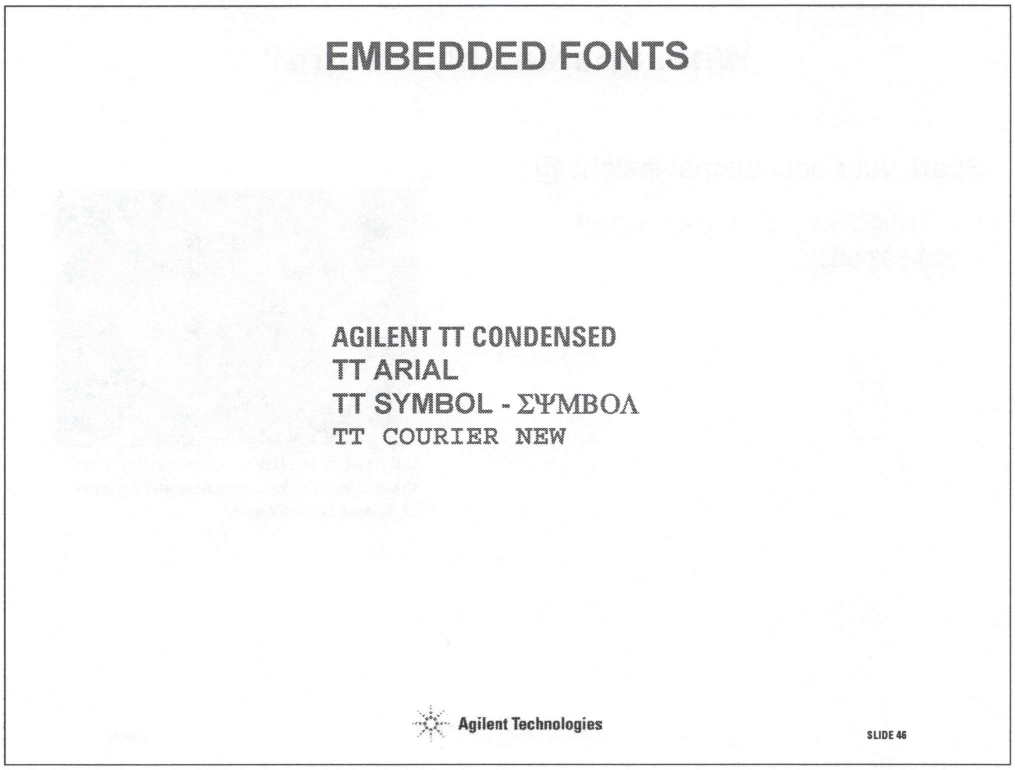
- www.agilent.com/find/tmeducation
- 800-593-6632



Bill Hewlett and Dave Packard founders of the Hewlett-Packard company and heritage of Agilent Technologies.

Agilent Technologies

SLIDE 45

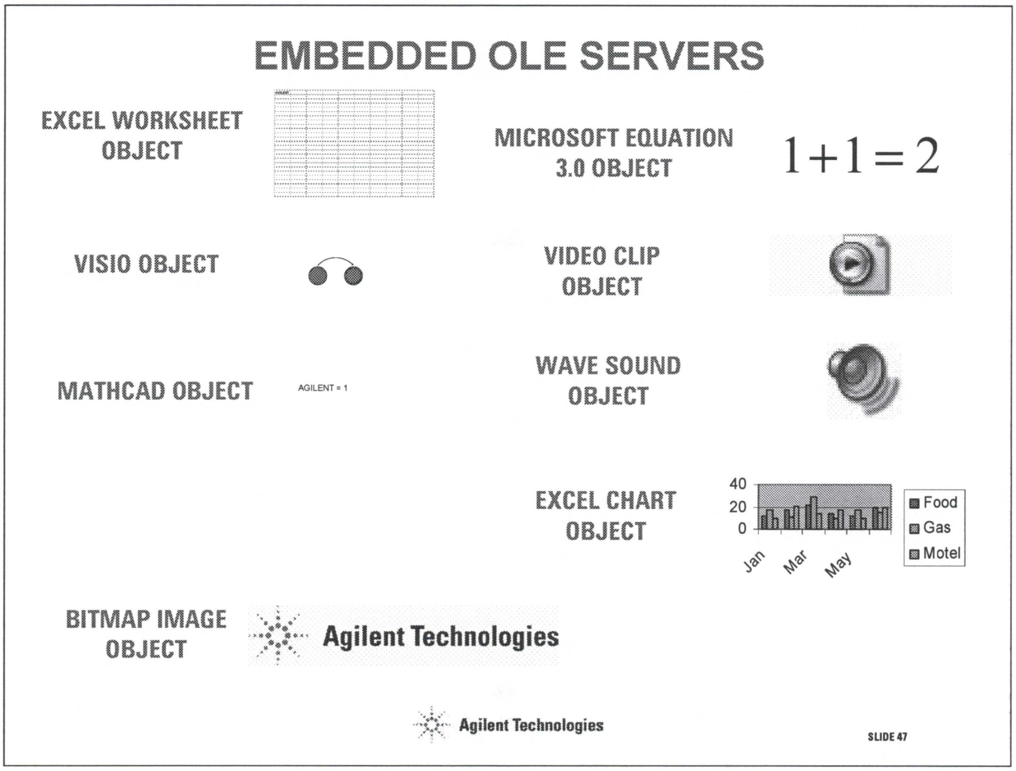


THIS SLIDE SHOWS THE EMBEDDED FONTS IN TEMPLATE.

TEMPLATE USERS ARE NOT ABSOLUTELY REQUIRED TO USE ONLY THE EMBEDDED FONTS, BUT IT IS RECOMMENDED.

IN PART, THE GENESIS OF THIS TEMPLATE, IS TO ELIMINATE THE FREQUENT PROBLEM OF SLIDES WITH FONTS THAT ARE NOT RESIDENT ON PC'S WITHIN THE USER COMMUNITY. BY LIMITING THE NUMBER OF FONTS AND EMBEDDING THEM INTO THE TEMPLATE, WE INTEND TO ELIMINATE THE PROBLEM OF SLIDE SETS THAT APPEAR CORRUPTED DUE TO LOCAL FONT SUBSTITUTION.

IF YOU REQUIRE AN ALTERNATIVE FONT FOR YOUR PRESENTATION, PLEASE EMBED THAT FONT IN YOUR PRESENTATION.



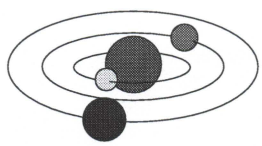
THIS SLIDE CONTAINS OLE OBJECTS AND CAUSES EMBEDDING OF THE RESPECTIVE OLE SERVERS IN THE TEMPLATE.

LEAVE THIS SLIDE AT THE VERY END OF THE PRESENTATION WHICH SHOULD FORCE THE OLE SERVERS TO REMAIN EMBEDDED IN YOUR PRESENTATION.

PLEASE EMBED ANY ADDITIONAL OLE SERVERS YOUR PRESENTATION MAY REQUIRE.

RF & Microwave Measurement Fundamentals

Digital Modulation



Agilent Technologies

Lesson Objectives

What is the advantage of digital signal?

How much frequency spectrum do digital signals need?

Reducing the bandwidth of the signal.

How complex modulations conserve bandwidth and ease some design aspects.

The signal space; constellation and vector diagrams How we view what's happening.

How we create these modulations.

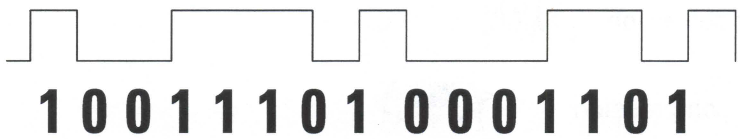
Noise and other degradations.

Agilent Technologies

SLIDE 2

Digital Signal - The Digital Baseband Signal

This is a non-return to zero (NRZ) signal

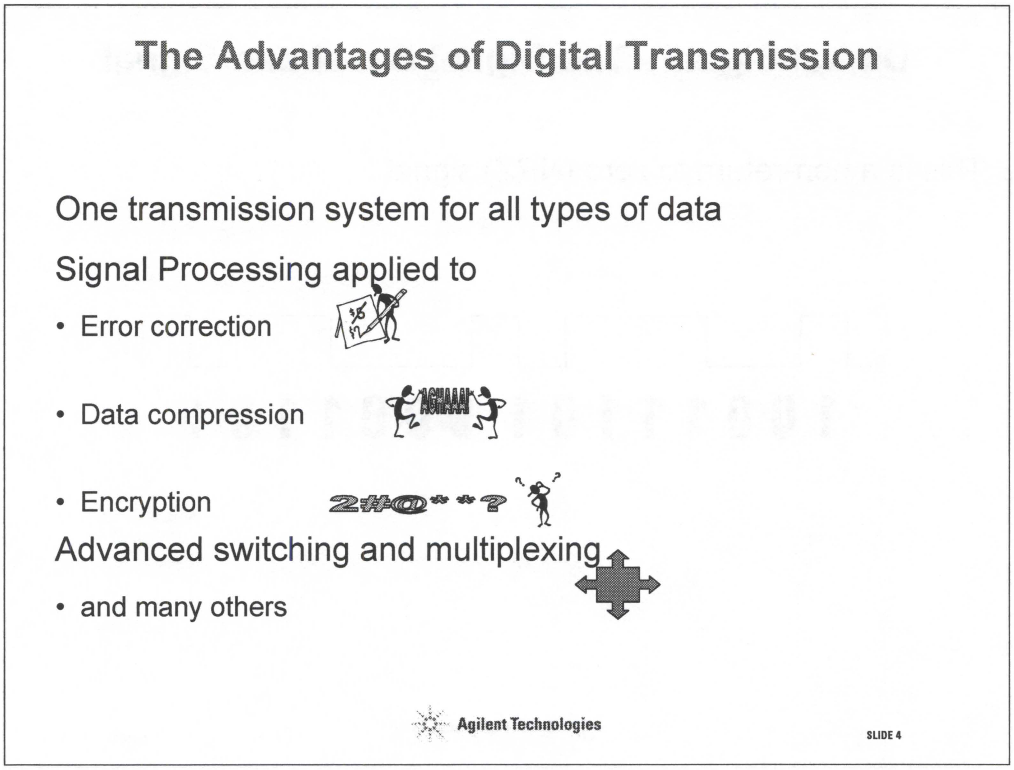


Agilent Technologies

SLIDE 3

The Digital Baseband Signal

This succession of one's and zero's may represent any data. It could be a fragment of music, speech, or a bank transaction. At this stage of the transmission our only concern is how to send it over long distances without significant data corruption. There are many levels of error checking which may be applied to the received data to ensure integrity of information.



In the past the coexistence of different program material such as video, telephone, and text on the same transmission medium was possible but each needed special handling. The ability to digitize material from all sources has unified the engineering solution, all sources are data!

With this comes other advantages such as the ability to secure sensitive material by various encryption schemes, to ensure data integrity by error correction and preprocessing the digitized signal to reduce the amount of data transmitted.

In modern communications networks the multiplexing and routing of messages can be handled by digital manipulation rather than by analog filters and switches.

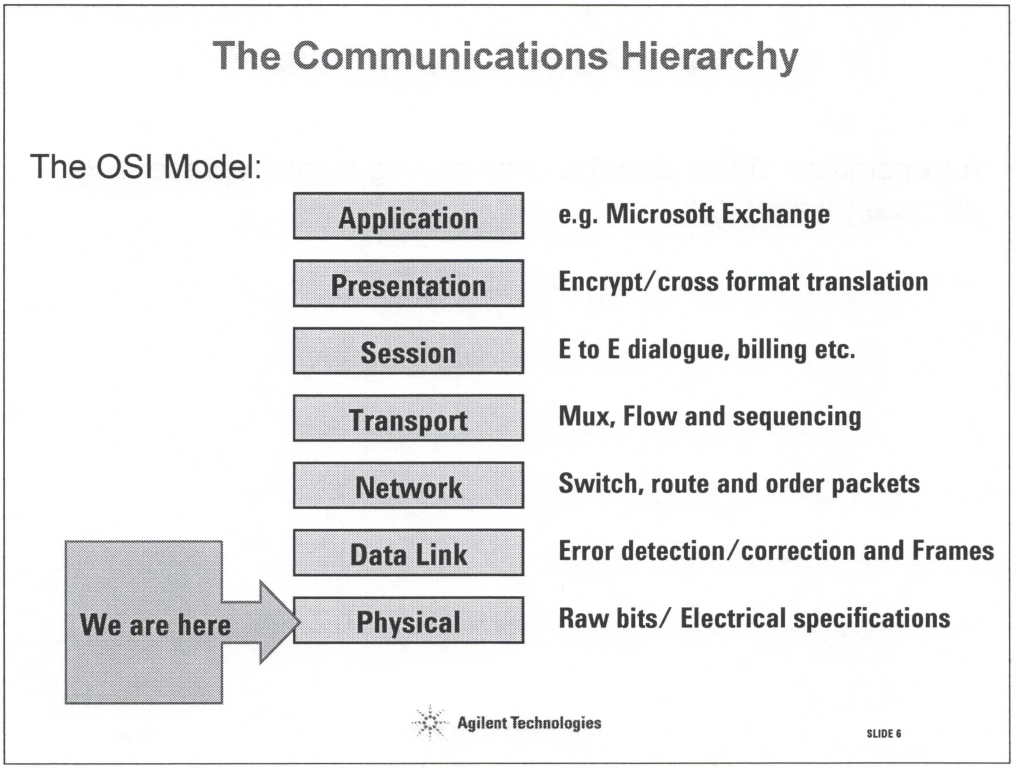
What is Digital Transmission?

A transmitted digital signal is is an analog signal representing discrete (digital) data.

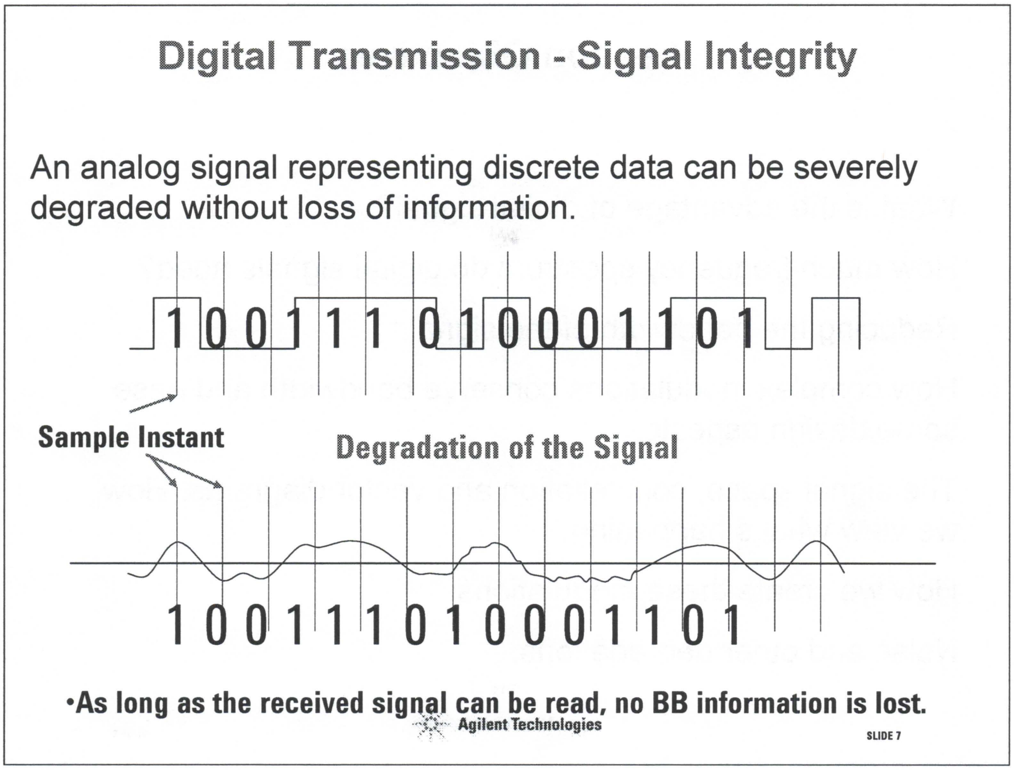
Agilent Technologies

SLIDE 5

The word “digital” implies discrete, however it is important to remember that any transmitted signal is analog and as such is subject to analog distortions. The data the signal represents is discrete, this data in turn may represent a discrete or analog application.



This is the standard model used in communications, its bottom layer is the Physical level and that is where we are in this lesson.



Bits, the elements of digital data, must be realized as an electrical signal, formatted in discrete values to represent these bits, or groups of bits. When a bit stream is presented to the transmission system it is has analog properties and as such is subject to distortion and interference. The only criteria for a successful transmission is that the received waveform is recognizable as a bit stream, in other words can a detector recreate the original message from this received waveform. This is the basic advantage of digital data transmission.

In the above picture we can see that there are two discrete quantities to be considered:

- Discrete times where the wave is sampled.
- Discrete levels which separate the logical values

In the above picture this is a binary signal and would need just one threshold, to separate a "one" from a "zero."

Lesson Objectives

What is the advantage of digital signal?

How much frequency spectrum do digital signals need?

Reducing the bandwidth of the signal.

How complex modulations conserve bandwidth and ease some design aspects.

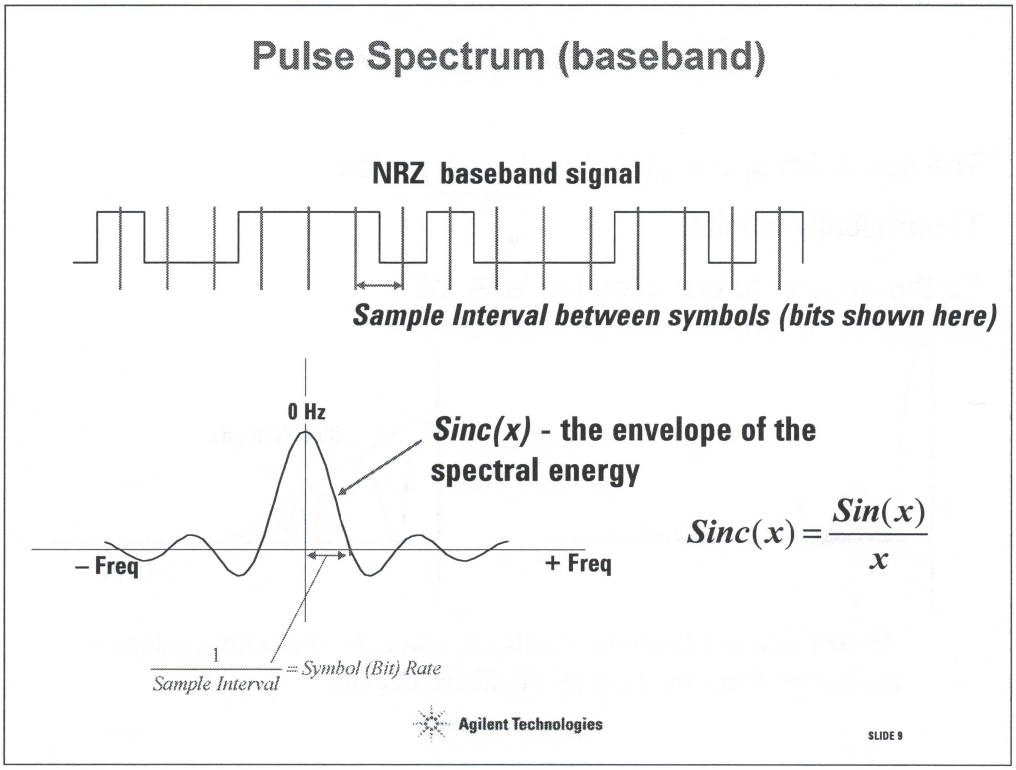
The signal space; constellation and vector diagrams. How we view what's happening.

How we create these modulations.

Noise and other degradations.

Agilent Technologies

SLIDE 8



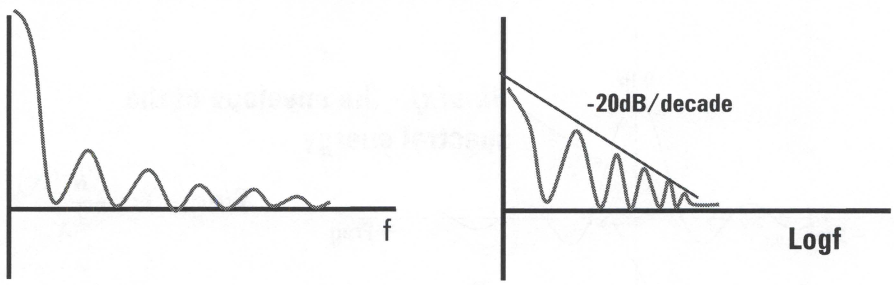
A binary signal if viewed in the frequency domain has wide spectrum which extends to ± infinity. The significance of the energy decreases as frequency increases, nevertheless a very large bandwidth would be needed to transmit the analog properties of this binary signal.

Transmission Bandwidth

The spectrum of a digital signal is very wide.

Theoretically infinite.

So the answer to our question is: A lot!



Roll Off rate is a function of pulse risetime, for 0 risetime pulses the power spectrum rolls off at –20dB/decade

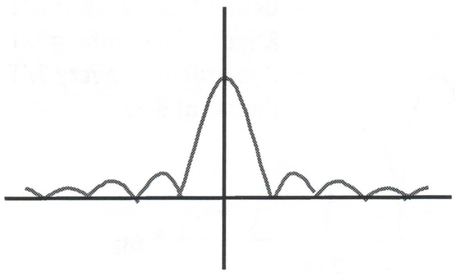
Agilent Technologies

SLIDE 10

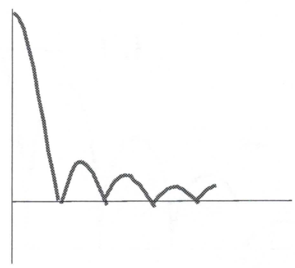
To send these pulses over a medium would need a very large bandwidth. The significant bandwidth can be reduced by controlling the risetime of the pulses, but as we shall see because all that is required of the receiver is to detect discrete levels at the sample instant, a much greater reduction in bandwidth is possible.

The Spectrum Analyzer View

Sinc(x) Magnitude



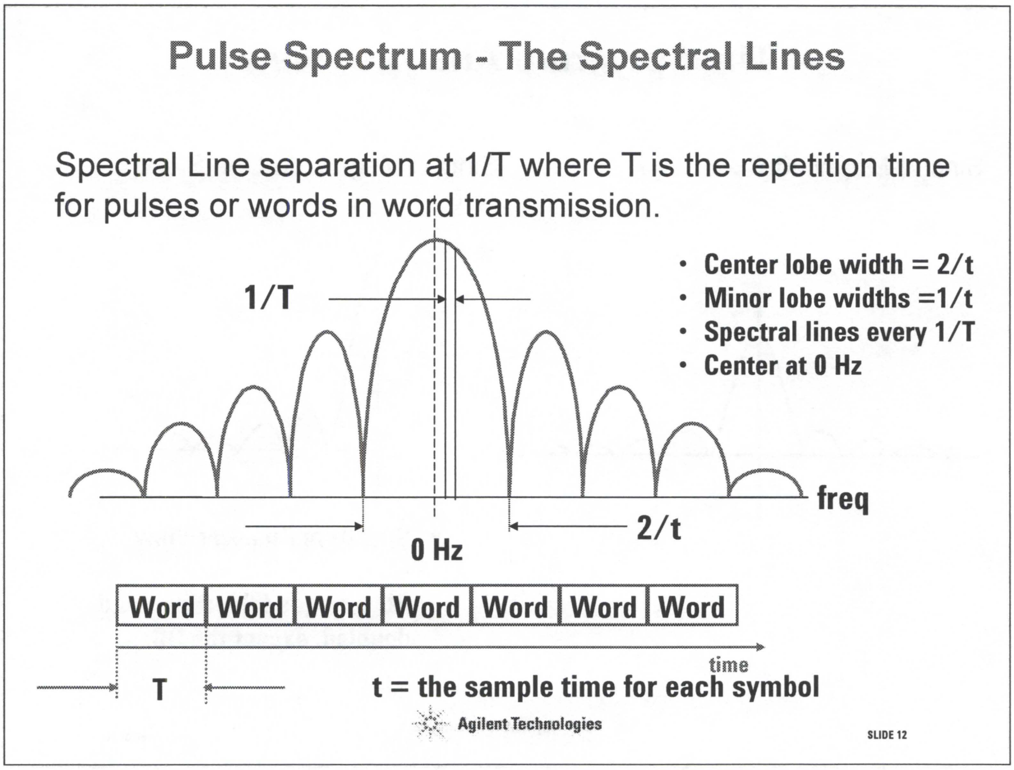
Sinc(x) Magnitude One Sided



- **Spectrum Analyzer View**
- **All voltages folded over and doubled, except the DC.**

Agilent Technologies

SLIDE 11


Pulse Spectrum - The Spectral Lines
Spectral Line separation at 1/T where T is the repetition time for pulses or words in word transmission.
1/T
• Center lobe width = 2/t
• Minor lobe widths =1/t
• Spectral lines every 1/T
• Center at 0 Hz
freq
2/t
0 Hz
Word
Word
Word
Word
Word
Word
Word
time
T
t = the sample time for each symbol
Agilent Technologies
SLIDE 12

Example:

A PRBS (2^7-1) NRZ signal is transmitted at 1MB/s

- What is the spectral line spacing?
- 2^7-1 = 127

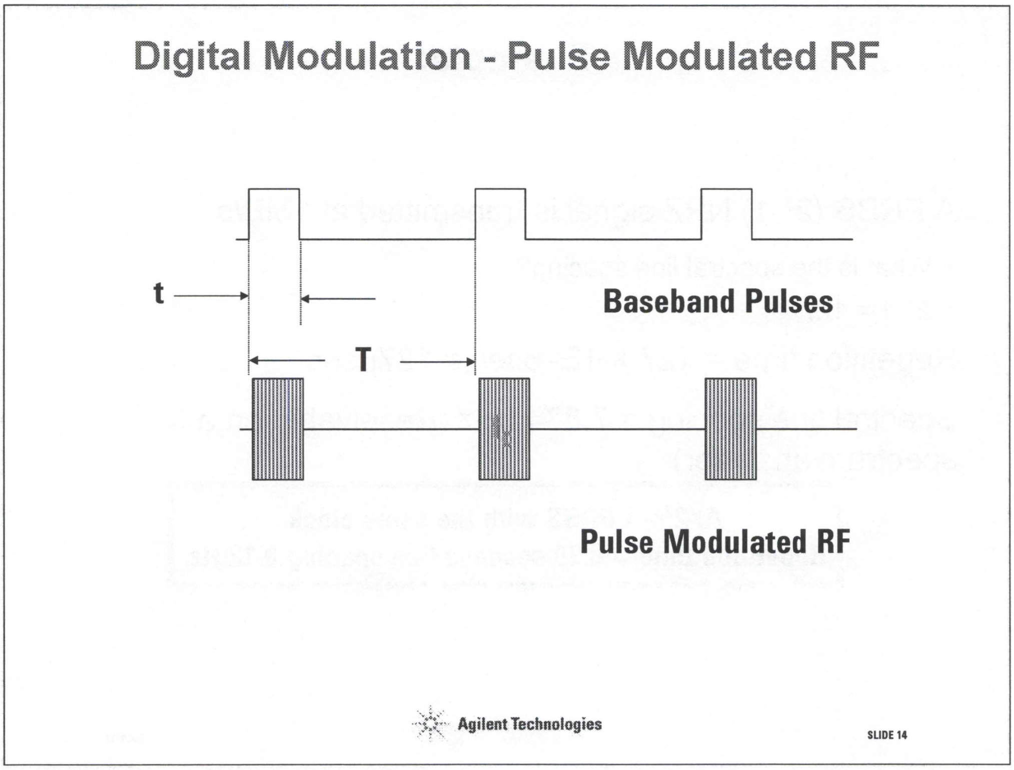
Repetition time = 127 × 1E–6sec = 127μsec

Spectral line spacing = 7.874 kHz (Resolvable on a spectrum analyzer)

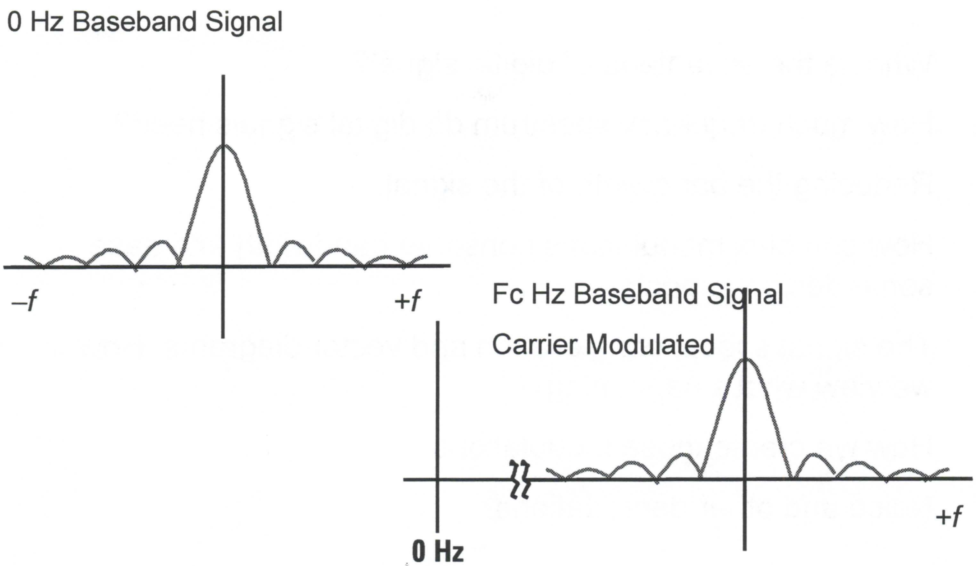
At 2^{23}–1 PRBS with the same clock
Repetition time = 8.39 seconds line spacing 0.12 Hz.

Agilent Technologies

SLIDE 13

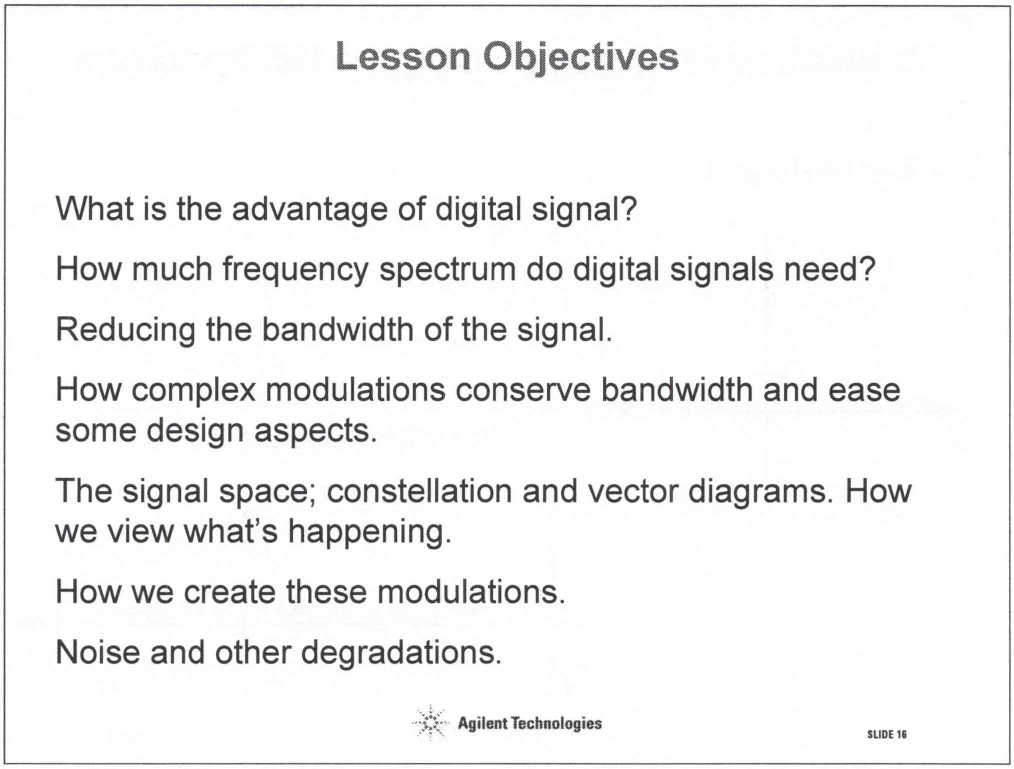

Digital Modulation - Pulse Modulated RF
t
T
Baseband Pulses
Pulse Modulated RF
Agilent Technologies
SLIDE 14

Modulation Frequency Shifts the BB Spectrum

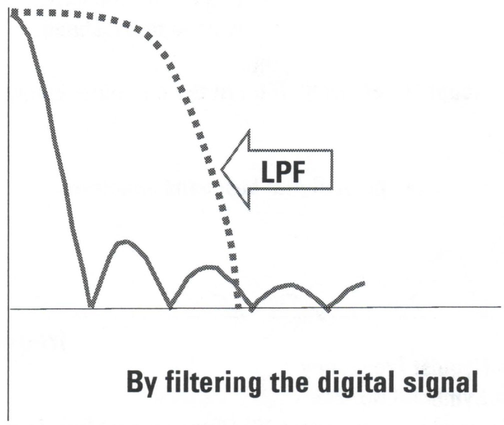


Agilent Technologies

SLIDE 15


Lesson Objectives
What is the advantage of digital signal?
How much frequency spectrum do digital signals need?
Reducing the bandwidth of the signal.
How complex modulations conserve bandwidth and ease some design aspects.
The signal space; constellation and vector diagrams. How we view what's happening.
How we create these modulations.
Noise and other degradations.
Agilent Technologies
SLIDE 16

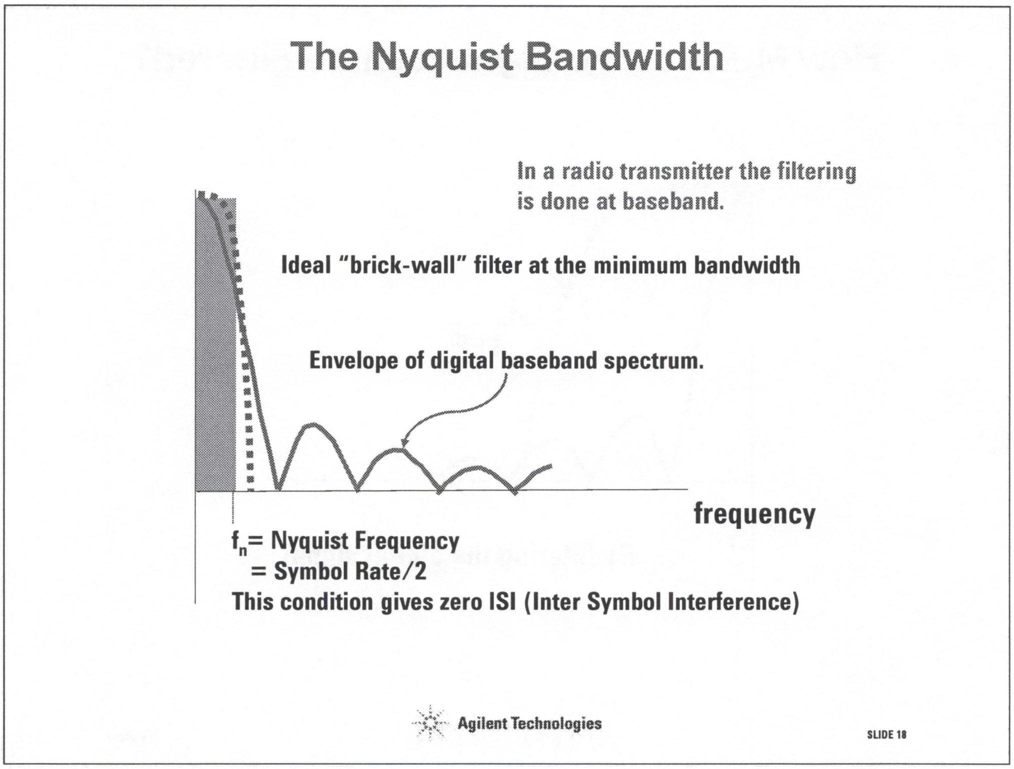
How Much of the Signal can be Filtered?



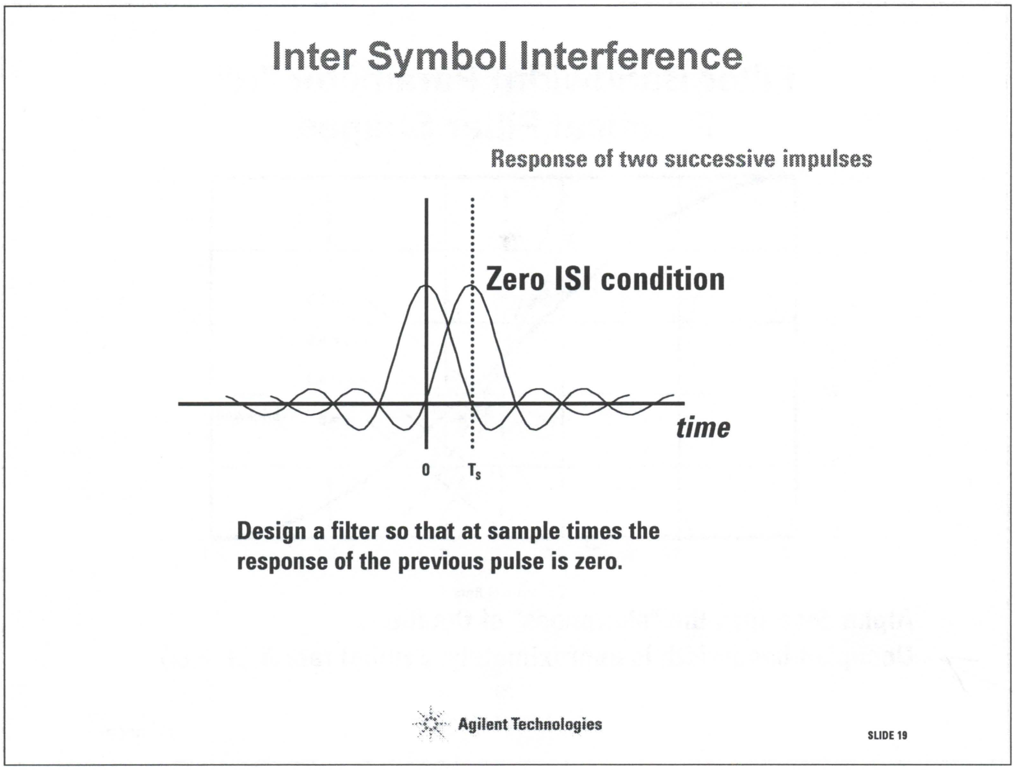
By filtering the digital signal

SLIDE 17

In a digital radio the filtering is done at baseband. If this BB signal is passed through a low pass filter we can reduce the bandwidth of the transmitted signal, the question is, by how much can this be done while still receiving a signal where the discrete levels are still recognizable?



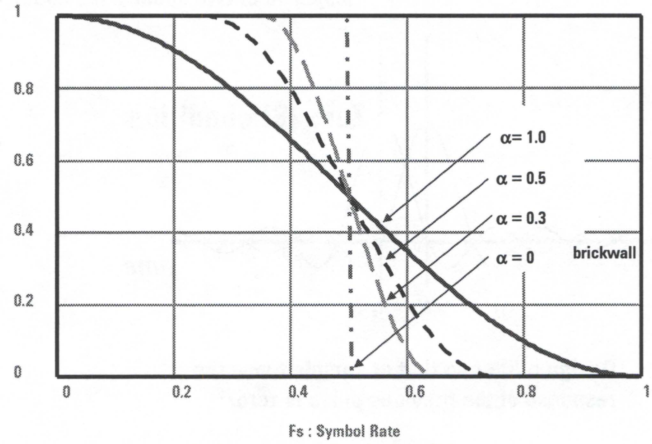
Nyquist, whose name often comes up in connection with communications electronics, analyzed the models of pulse transmission an concluded that a perfect low pass filter bandwidth fs/2 could transmit such that successive pulses would not interfere with levels at the sampling instant.



This Nyquist filter will "ring", but at exactly fs/2 the ringing is such that adjacent pulses do not interfere at the sampling instant. This is called the zero ISI condition. ISI = Inter Symbol Interference.

This perfect "brickwall filter" is not realizable and even if it was the bandwidth would have to be exact; no tolerance!

Filter Bandwidth Parameter "α"
Practical Filter Shapes



Alpha describes the "sharpness" of the filter.
Occupied bandwidth is approximately: Symbol rate X (1 + α)

Agilent Technologies

SLIDE 20

The Nyquist condition (α=0) is a theoretical but useful limit. It provides the basis for the design of the practical filters shown above. A typical α is 0.3, this shapes the sinc spectrum but extends the bandwidth beyond the fs/2 limit.

The occupied bandwidth is now (1 + α)×Symbol rate

Lesson Objectives

What is the advantage of digital signal?

How much frequency spectrum do digital signals need?

Reducing the bandwidth of the signal.

How complex modulations conserve bandwidth and ease some design aspects.

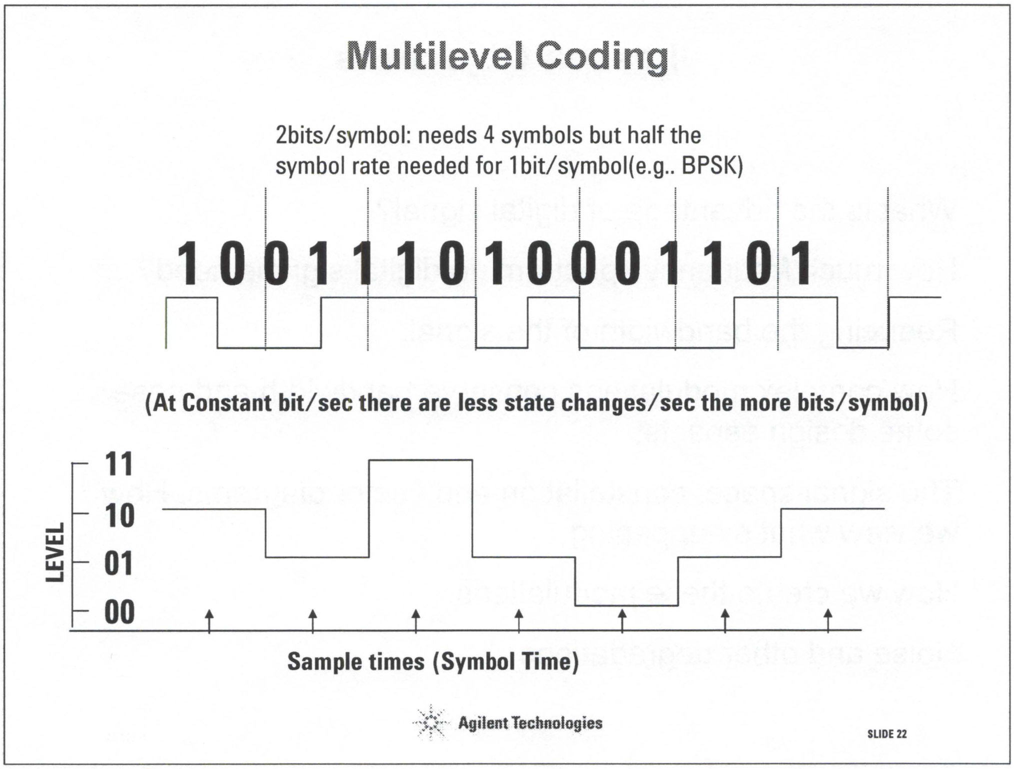
The signal space; constellation and vector diagrams. How we view what's happening.

How we create these modulations.

Noise and other degradations.

Agilent Technologies

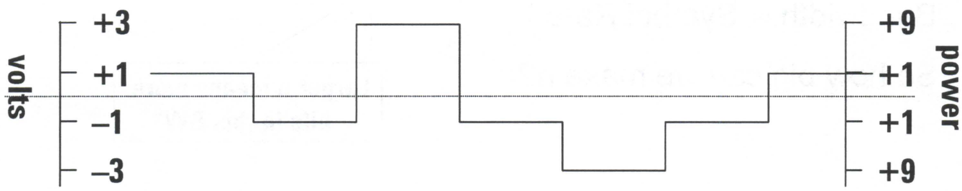
SLIDE 21


Multilevel Coding
2bits/symbol: needs 4 symbols but half the
symbol rate needed for 1bit/symbol(e.g.. BPSK)
1 0 0 1 1 1 0 1 0 0 0 1 1 0 1
(At Constant bit/sec there are less state changes/sec the more bits/symbol)
11
10
01
00
LEVEL
Sample times (Symbol Time)
Agilent Technologies
SLIDE 22

Multilevel Coding and Power

If amplitude coding were used for multilevel transmission the power requirements increase compared to binary.

If all symbols were equally probable then in the case of 4 levels 5 times (7dB) more power is needed to get same S/N.



Agilent Technologies

SLIDE 23

There are, as we shall see later, bandwidth economies to be made by creating a multilevel signal from the binary bit stream. In this picture, the binary word is grouped by twos. Since there are four different ways to combine two bits the new signal has four levels. It can also be seen in this case that the sample rate is half the bit rate. The groupings are now called symbols and the sample rate the is the symbol rate.

The levels may be represented by amplitude, phase, frequency or any measurable quality of the received signal.

These groupings may be greater than two bits for example three bit symbols would need eight levels and four bit symbols would need sixteen levels etc.

In general: If n = number of bits per symbol, and m = the number of levels then:

$$m = 2^n$$

Symbol Rate

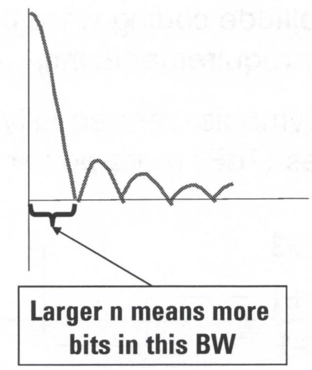
Number of symbols = 2^n

n = bits per symbol

Symbol rate = $(1/n) \times$ bit rate

Bandwidth $\propto$ Symbol Rate

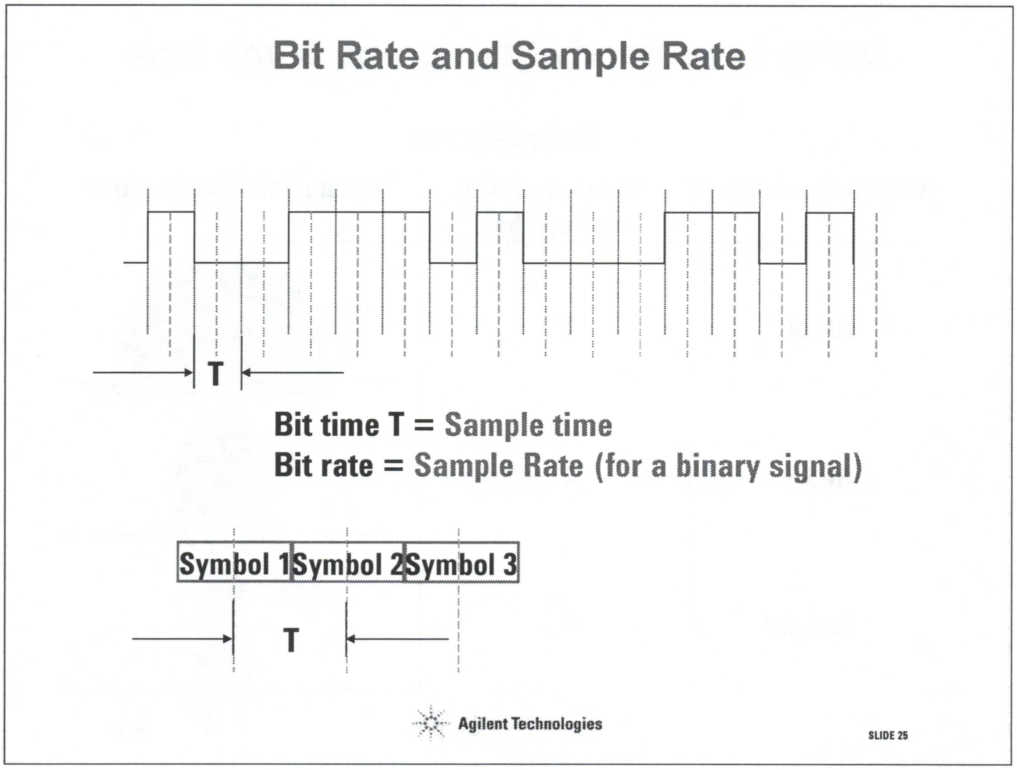
So how big can we make n?



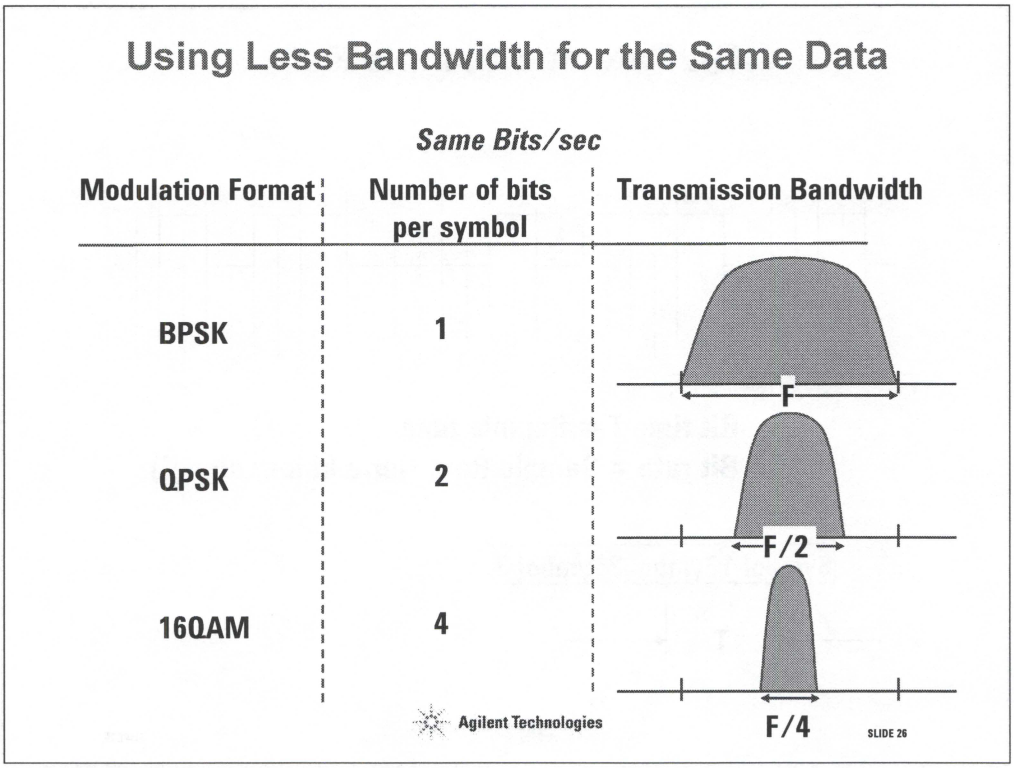
Agilent Technologies

SLIDE 24

Bandwidth is proportional to symbol rate, not the bit rate, by grouping bits into symbols we can get more bandwidth efficient.



For a binary signal the sample rate and bit rate are the same.



Here are comparisons of three modulation formats, transferring information at the same rate. The meaning of the formats will be explained later

Other Discrete (Digital) Modulations

Pulse (ON/OFF)

Phase (Discrete phases)

Amplitude(Discrete Amplitudes)

Frequency (FSK)

Combinations of Phase and Amplitude

Agilent Technologies

SLIDE 27

Lesson Objectives

What is the advantage of digital signal?

How much frequency spectrum do digital signals need?

Reducing the bandwidth of the signal.

The signal space; constellation and vector diagrams. How we view what's happening.

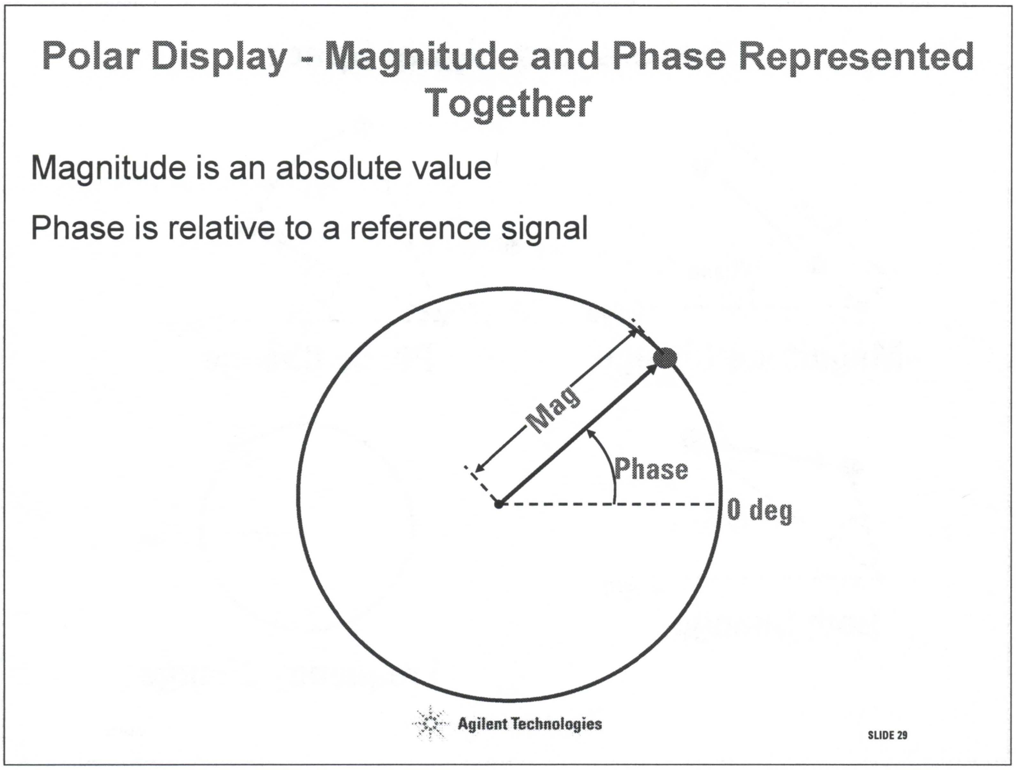
How complex modulations conserve bandwidth and ease some design aspects.

How we create these modulations.

Noise and other degradations.

Agilent Technologies

SLIDE 28



This model of a sinusoid is very commonly used, instead of the amplitude vs. angle "wavy line" graph, it can be more insightful to use the vector view where magnitude is the length of the vector and the angle(phase) is the angle made with a fixed datum. If the above represented $v(t) = A \cos(\omega t)$ then the length of the vector would be A and the angle ωt. This angle starts at t=0 at 0degrees,as time progresses the angle would increase at a constant rate, so the vector would rotate counter clockwise at a constant speed.

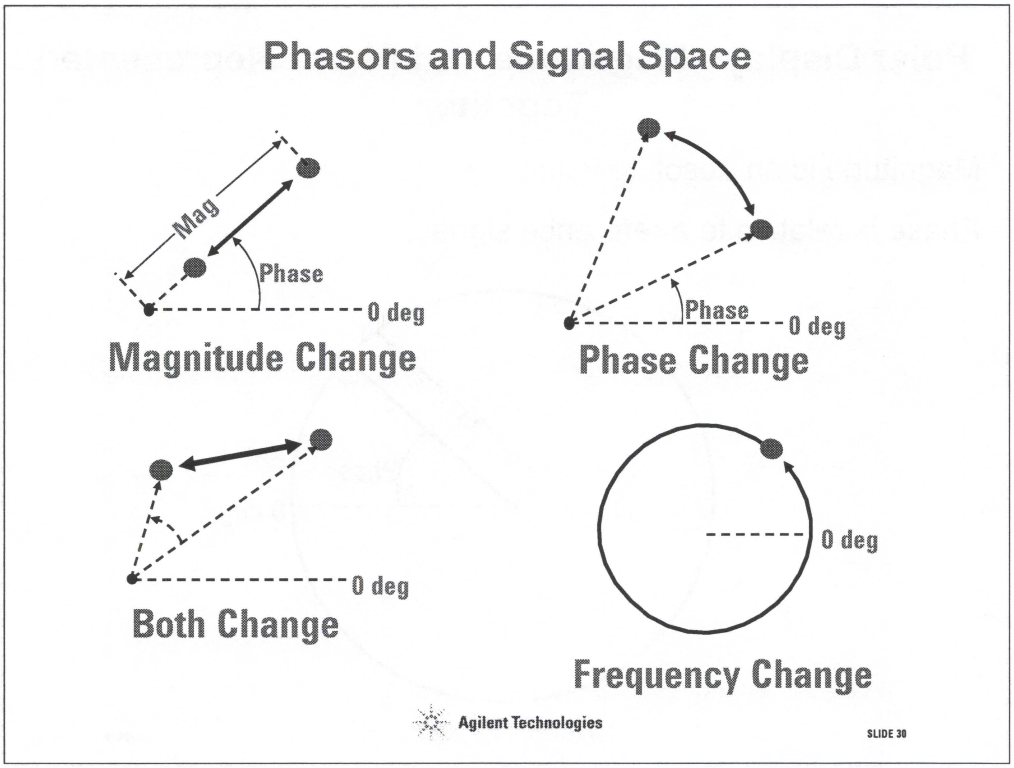
If we, as an observer, rotate at the same speed then the vector will appear as a stationary vector. Thus any changes relative to this will be seen relative to this reference.

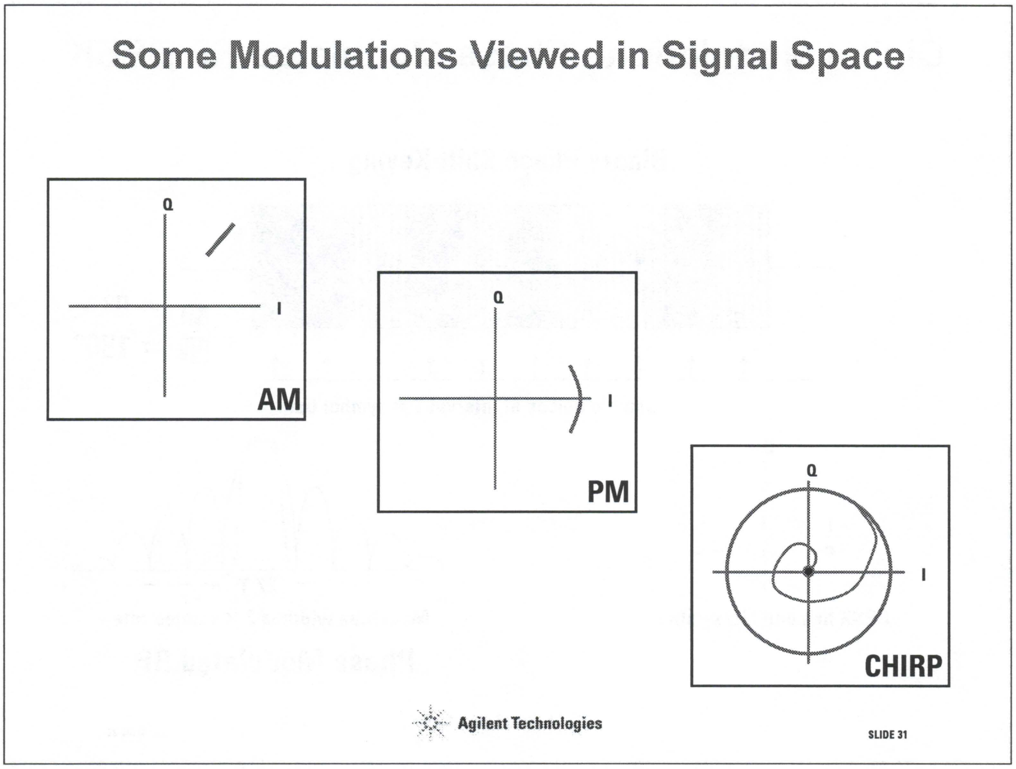
Signal Space

As the signal changes with respect to the reference, the tip of the vector will move in this plane we call this plane the *signal space*.

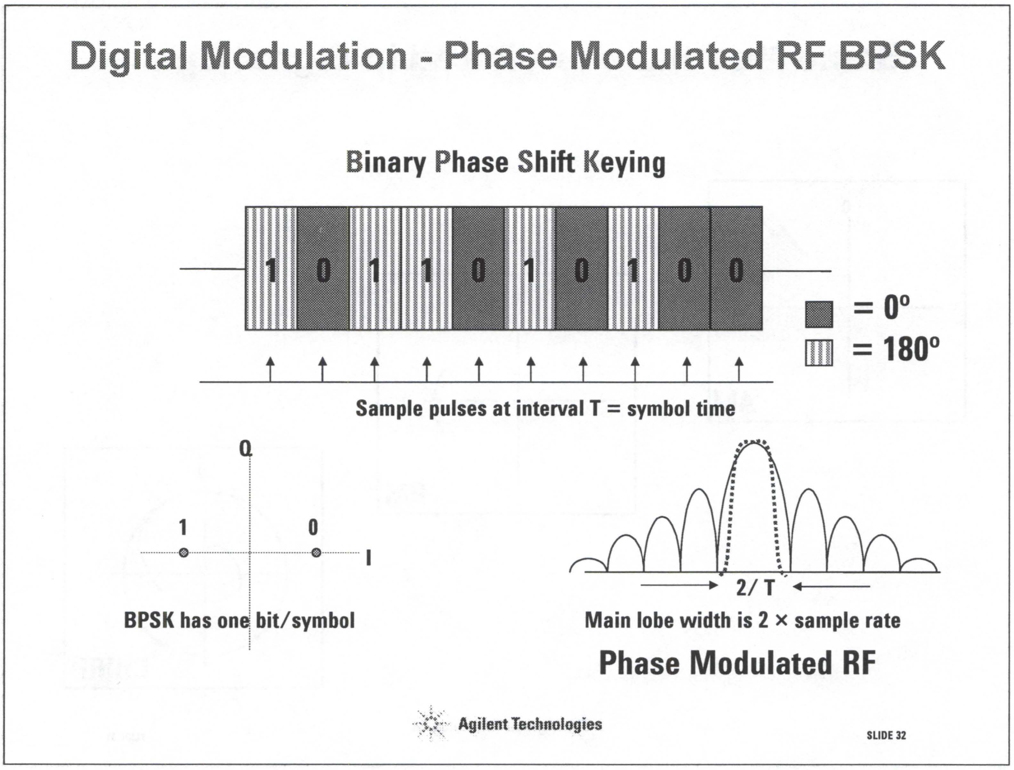
Phasors and Vectors

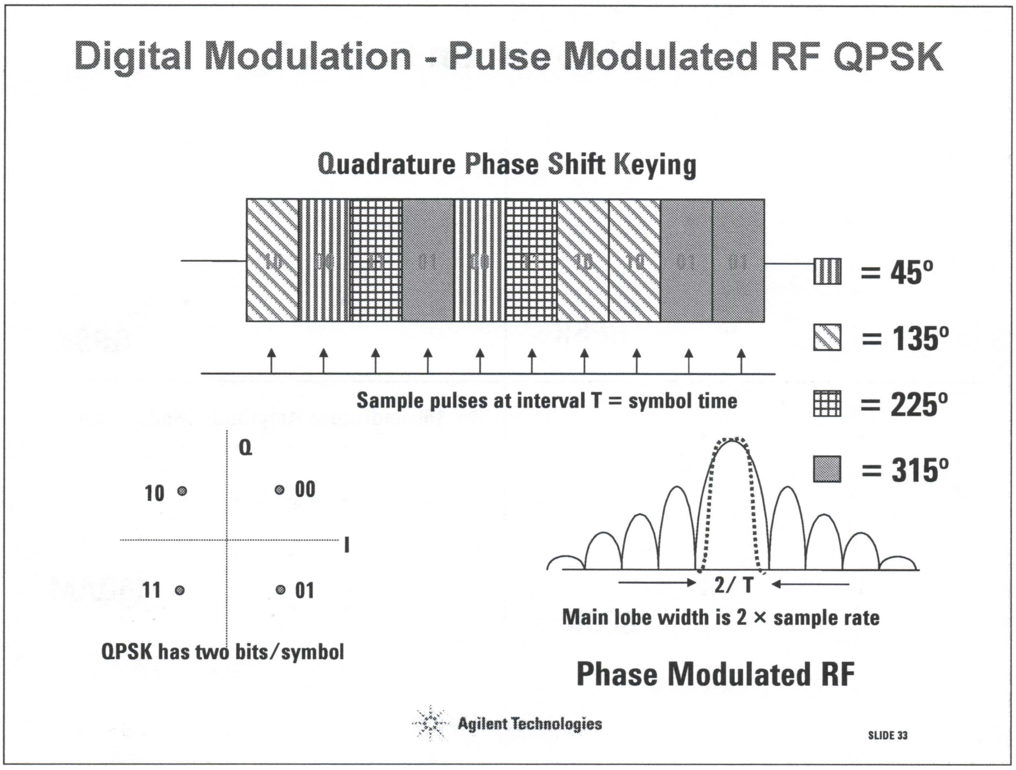
This vector which is anchored at one end has a special name, this is called a *phasor*.

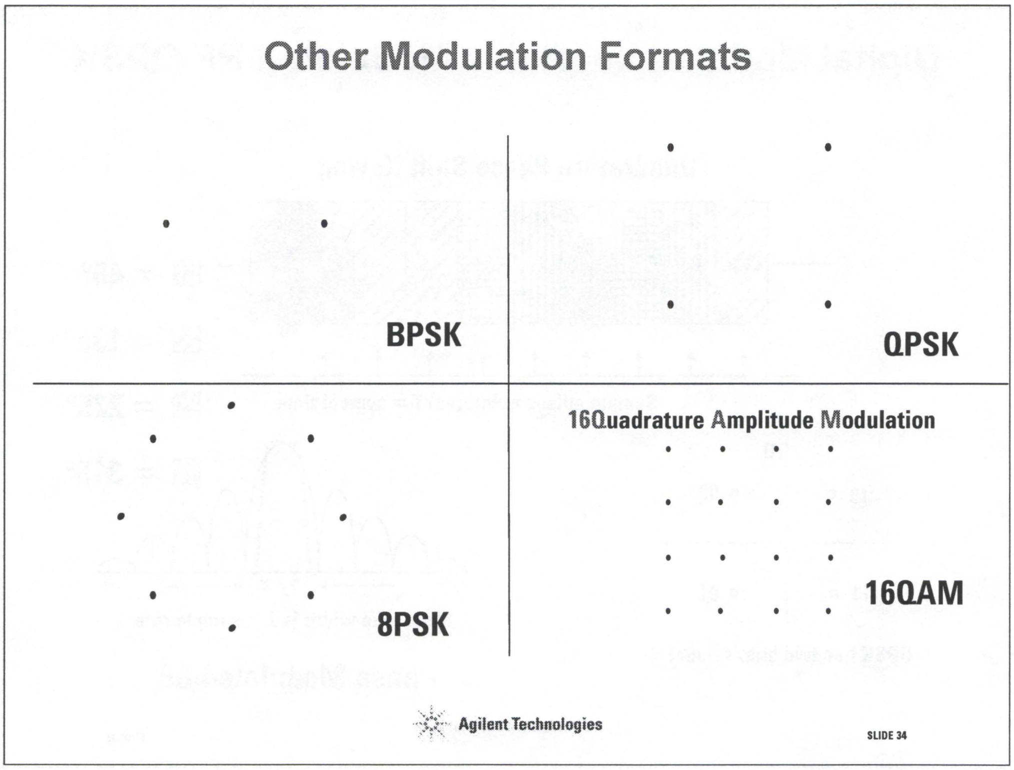

Phasors and Signal Space
Mag
Phase
0 deg
Magnitude Change
Phase
0 deg
Phase Change
0 deg
Both Change
0 deg
Frequency Change
Agilent Technologies
SLIDE 30



The trajectories of some phasors are shown, they can give us detailed information, for example the chirp, which is a frequency sweep from low to high, seems to have a switch on transient where the signal grows in amplitude and is increasing in frequency with respect to the reference.


Digital Modulation - Phase Modulated RF BPSK
Binary Phase Shift Keying
1 0 1 1 0 1 0 1 0 0
= 0°
= 180°
Sample pulses at interval T = symbol time
Q
1
0
I
BPSK has one bit/symbol
2/ T
Main lobe width is 2 × sample rate
Phase Modulated RF
Agilent Technologies
SLIDE 32


Digital Modulation - Pulse Modulated RF QPSK
Quadrature Phase Shift Keying
= 45°
= 135°
= 225°
= 315°
Sample pulses at interval T = symbol time
Q
10
00
I
11
01
QPSK has two bits/symbol
2/ T
Main lobe width is 2 × sample rate
Phase Modulated RF
Agilent Technologies
SLIDE 33


Other Modulation Formats
BPSK
QPSK
16Quadrature Amplitude Modulation
8PSK
16QAM
Agilent Technologies
SLIDE 34

Lesson Objectives

What is the advantage of digital signal?

How much frequency spectrum do digital signals need?

Reducing the bandwidth of the signal.

How complex modulations conserve bandwidth and ease some design aspects.

The signal space; constellation and vector diagrams. How we view what's happening.

How we create these modulations.

Noise and other degradations.

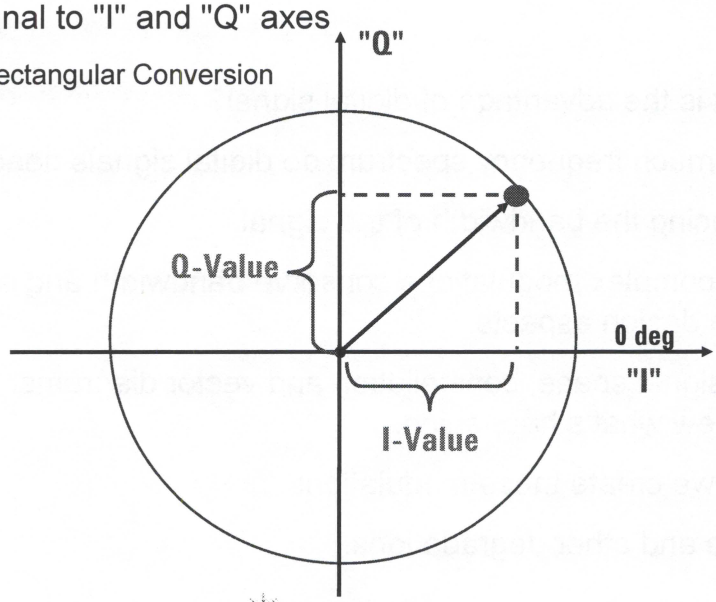
Agilent Technologies

SLIDE 35

Polar vs. "I-Q" Format

Project signal to "I" and "Q" axes

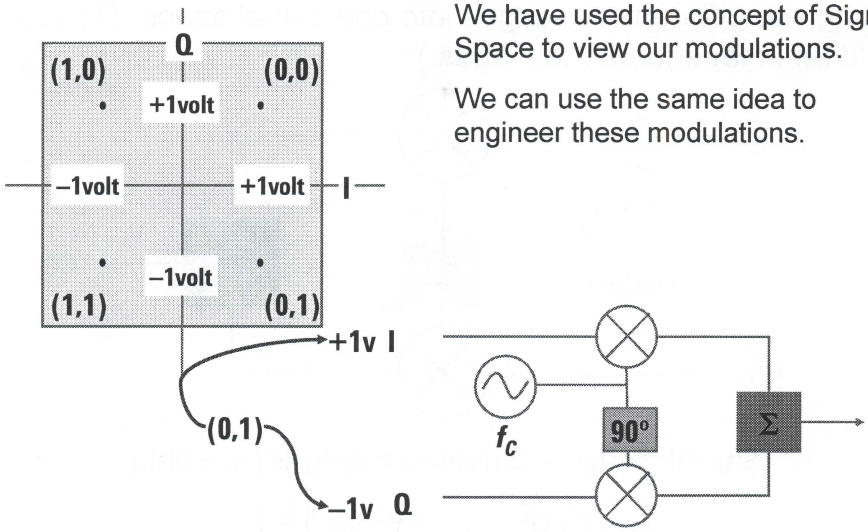
- Polar to Rectangular Conversion



Agilent Technologies

SLIDE 36

Creating Digital Modulation



We have used the concept of Signal Space to view our modulations.

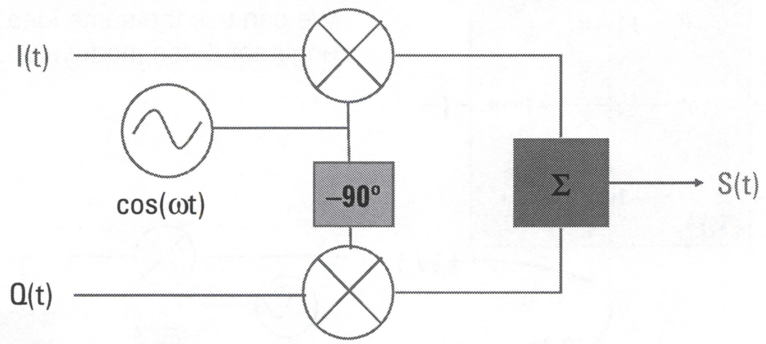
We can use the same idea to engineer these modulations.

Agilent Technologies

SLIDE 37

How Does this Work?

Putting two different messages into one signal space. (These could be independent messages.)



$$S(t) = I(t)\cos(\omega t) - Q(t)\sin(\omega t) = A(t)[\cos(\omega t + \theta(t))]$$

$$A = \sqrt{I^2 + Q^2} \qquad \theta = \tan^{-1}\left(\frac{Q}{I}\right)$$

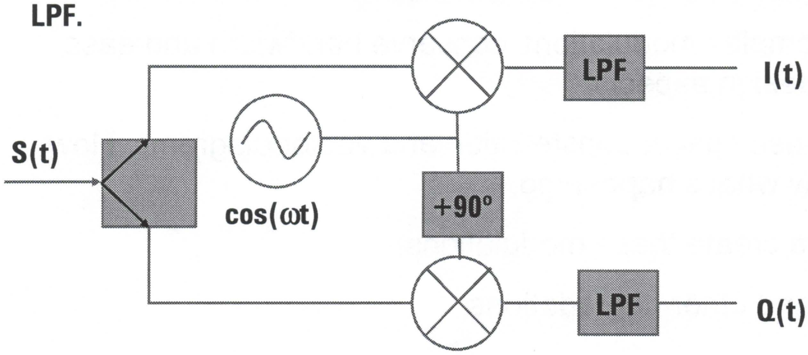
Agilent Technologies

SLIDE 38

Separating the Components: The Receiver

$$S(t) = I(t)\cos(\omega t) - Q(t)\sin(\omega t)$$

- **The composite signal is separated by multiplying (mixing) by sin(ωt) and cos (ωt), the resulting $\sin^2$ and $\cos^2$ terms become [I(t) or Q(t)] × ½[1 ± cos(2ωt)] terms - the 2ωt's are removed by the LPF.**



Agilent Technologies

SLIDE 39

Lesson Objectives

What is the advantage of digital signal?

How much frequency spectrum do digital signals need?

Reducing the bandwidth of the signal.

How complex modulations conserve bandwidth and ease some design aspects.

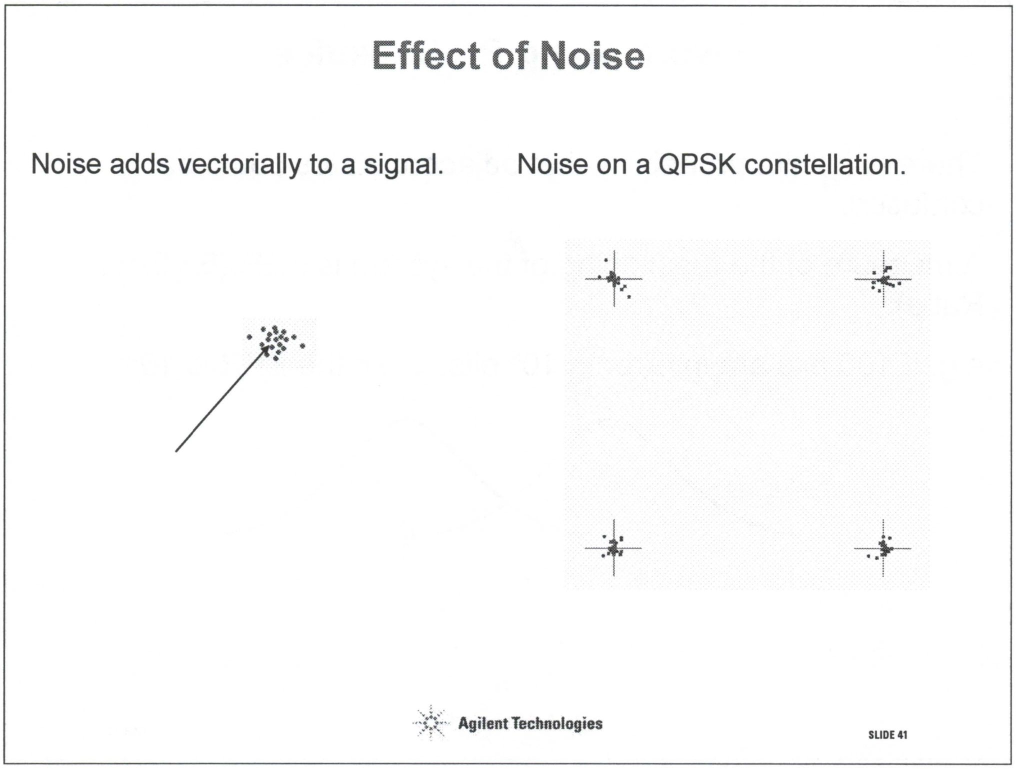
The phase space; constellation and vector diagrams. How we view what's happening.

How we create these modulations.

Noise and other degradations.

Agilent Technologies

SLIDE 40



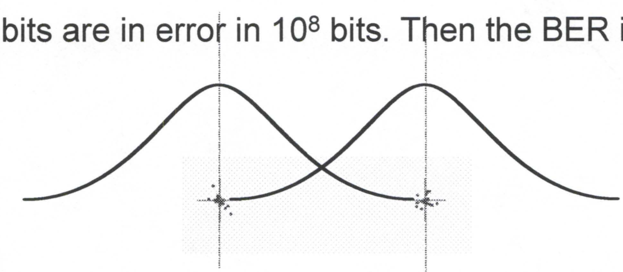
Noise adds vectorially to the signal which can be easily seen in the signal space.

Overlapping Probabilities

There is a finite probability that adjacent states could be confused.

A measure of the functioning of the system is BER (Bit Error Ratio)

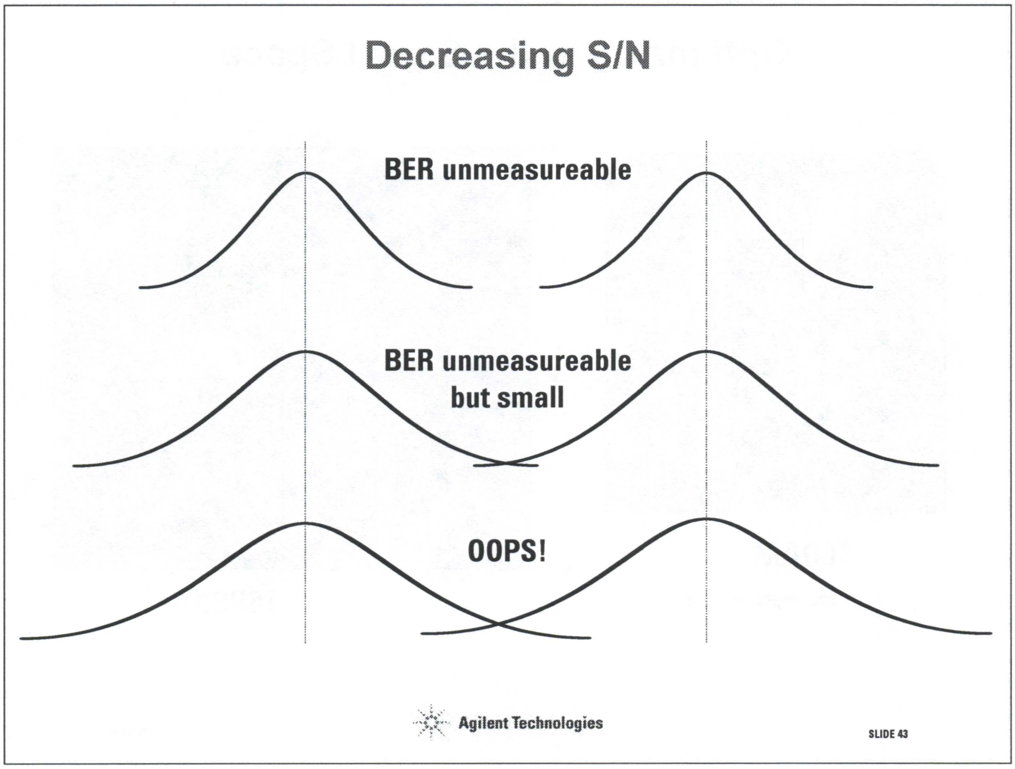
e.g. if 100 bits are in error in 10^8 bits. Then the BER is 10^{-6}



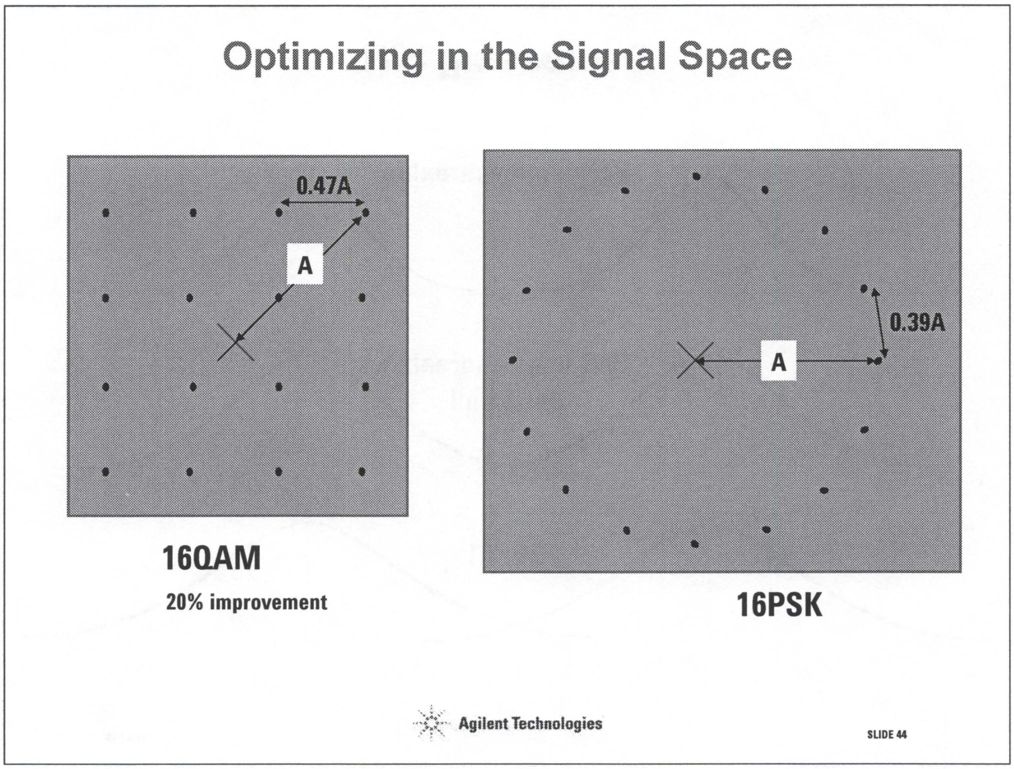
Agilent Technologies

SLIDE 42

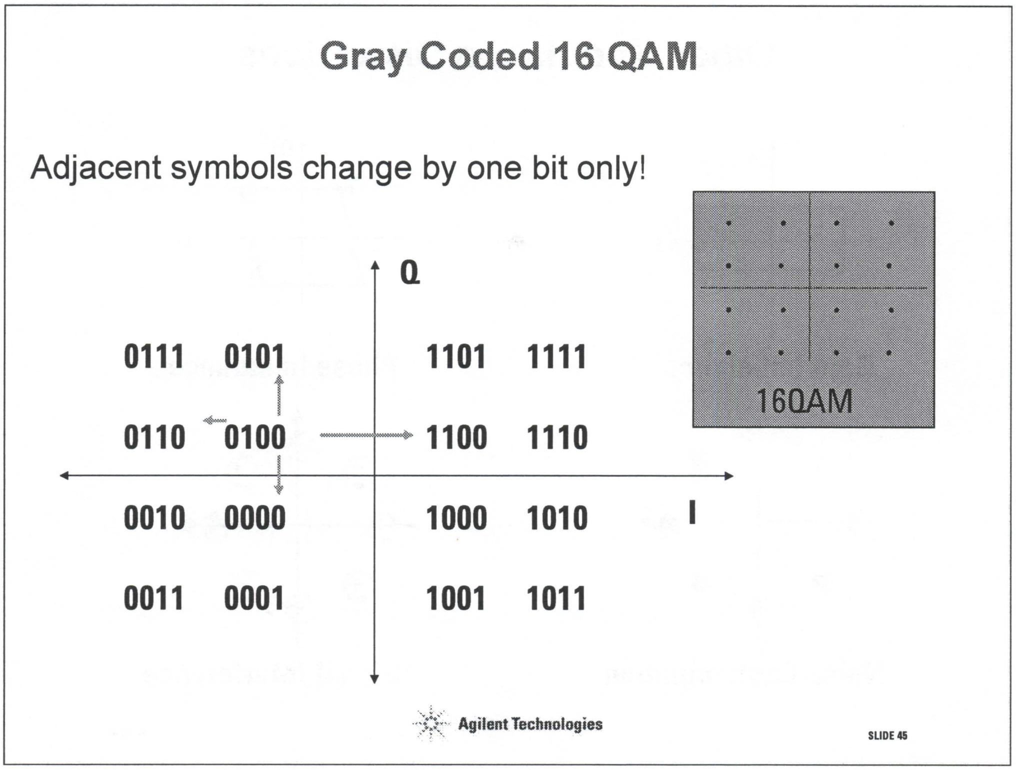
The noise distributions are mainly Gaussian with overlapping tails in the signal space. This means there is a small but finite probability of detection error.



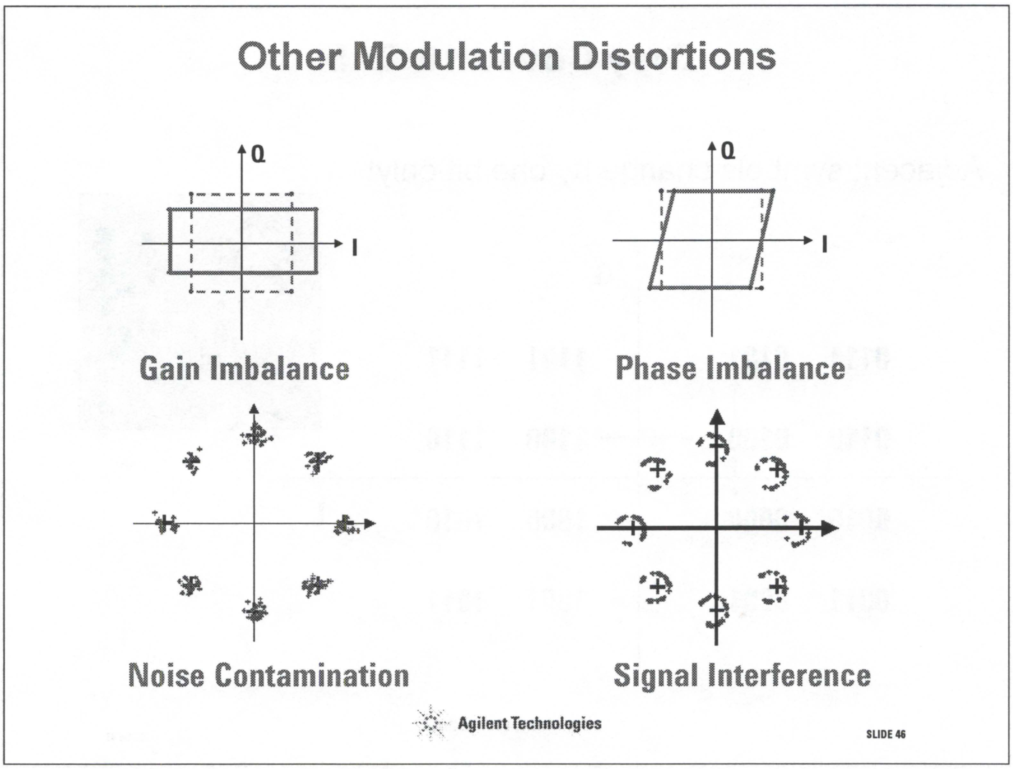
As the signal to noise ratio decreases the tails overlap even more making P_e(Probably of error) larger. In the above picture the mean value is right on the sample point but the standard deviation of the noise distribution is increasing until P_e is easily measurable.

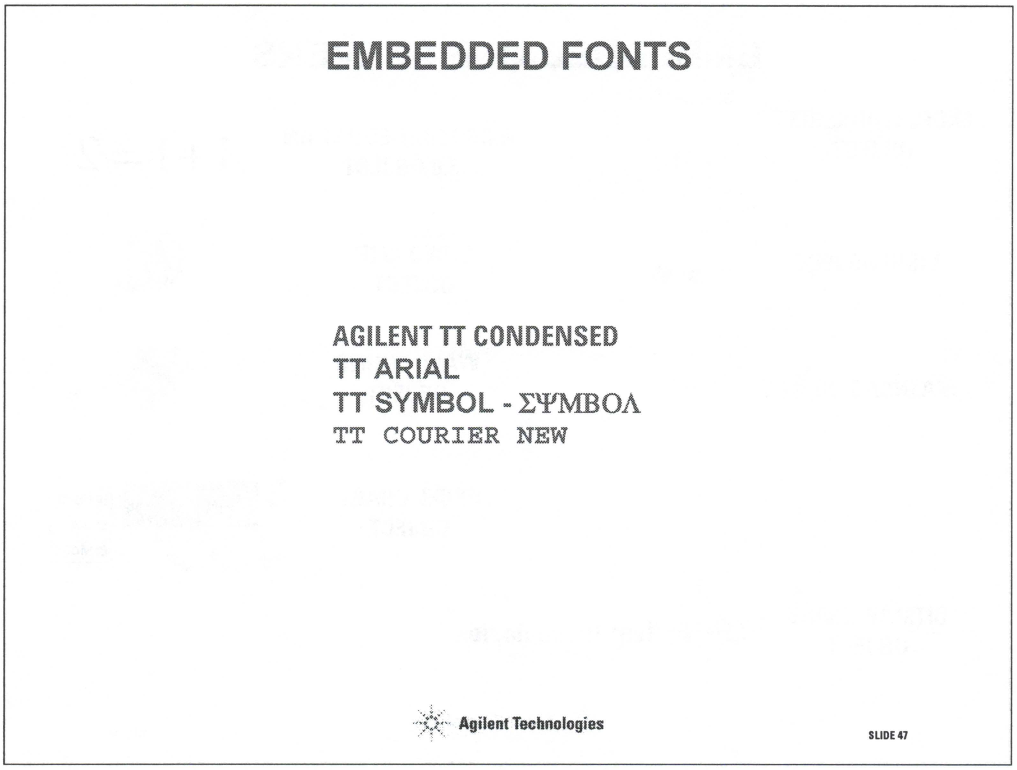


Another method to decrease the probability of error is to change use the signal space with a more even distribution of sample points. In this example the improvement of 16QAM to 16PSK is noted.



The way in which the symbols are mapped to the signal space is not usually arbitrary. A special arrangement Gray coding where adjacent states only differ by one bit is usually used. If the receiver interprets a state as an adjacent state only one bit will be in error.


Other Modulation Distortions
Q
I
Gain Imbalance
Q
I
Phase Imbalance
Noise Contamination
Signal Interference
Agilent Technologies
SLIDE 46

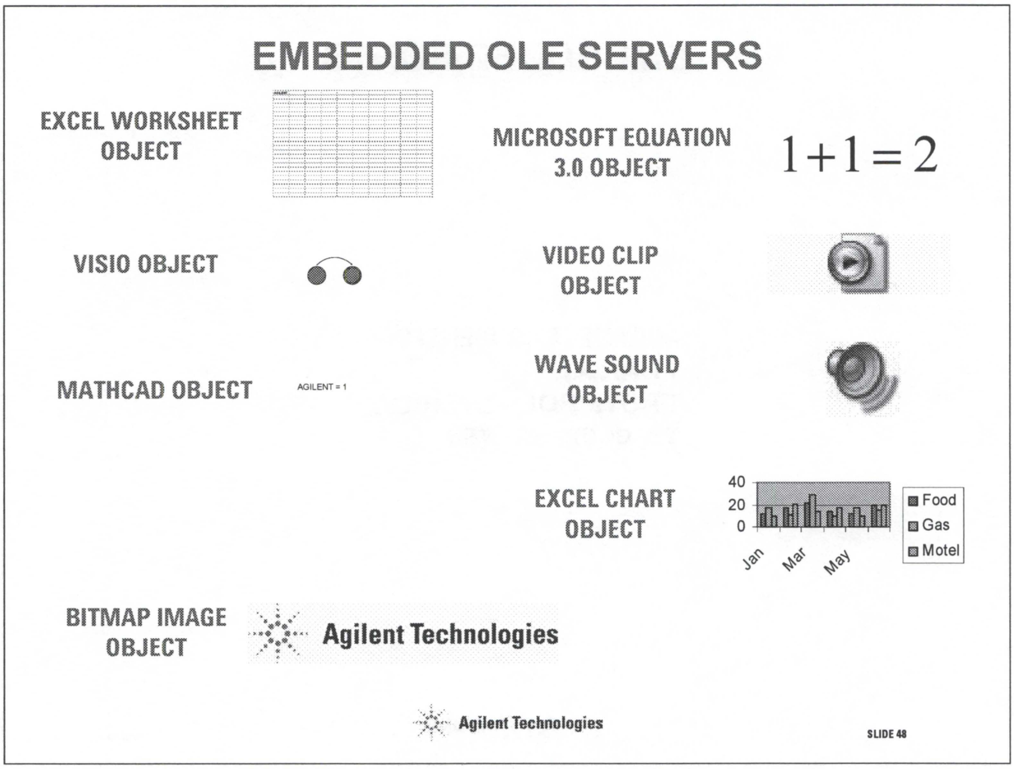


THIS SLIDE SHOWS THE EMBEDDED FONTS IN TEMPLATE.

TEMPLATE USERS ARE NOT ABSOLUTELY REQUIRED TO USE ONLY THE EMBEDDED FONTS, BUT IT IS RECOMMENDED.

IN PART, THE GENESIS OF THIS TEMPLATE, IS TO ELIMINATE THE FREQUENT PROBLEM OF SLIDES WITH FONTS THAT ARE NOT RESIDENT ON PC'S WITHIN THE USER COMMUNITY. BY LIMITING THE NUMBER OF FONTS AND EMBEDDING THEM INTO THE TEMPLATE, WE INTEND TO ELIMINATE THE PROBLEM OF SLIDE SETS THAT APPEAR CORRUPTED DUE TO LOCAL FONT SUBSTITUTION.

IF YOU REQUIRE AN ALTERNATIVE FONT FOR YOUR PRESENTATION, PLEASE EMBED THAT FONT IN YOUR PRESENTATION.



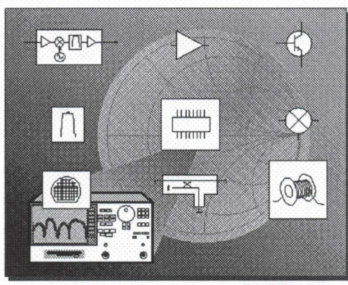
THIS SLIDE CONTAINS OLE OBJECTS AND CAUSES EMBEDDING OF THE RESPECTIVE OLE SERVERS IN THE TEMPLATE.

LEAVE THIS SLIDE AT THE VERY END OF THE PRESENTATION WHICH SHOULD FORCE THE OLE SERVERS TO REMAIN EMBEDDED IN YOUR PRESENTATION.

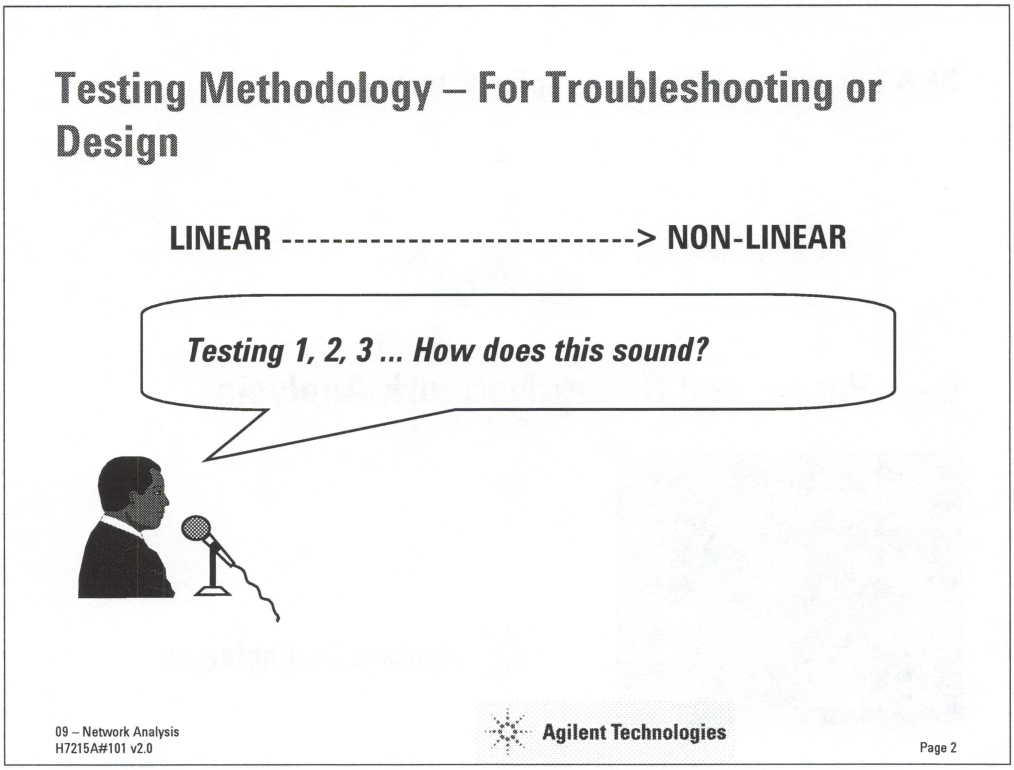
PLEASE EMBED ANY ADDITIONAL OLE SERVERS YOUR PRESENTATION MAY REQUIRE.

RF & Microwave Measurement Fundamentals

9 — Vector and Scalar Network Analysis

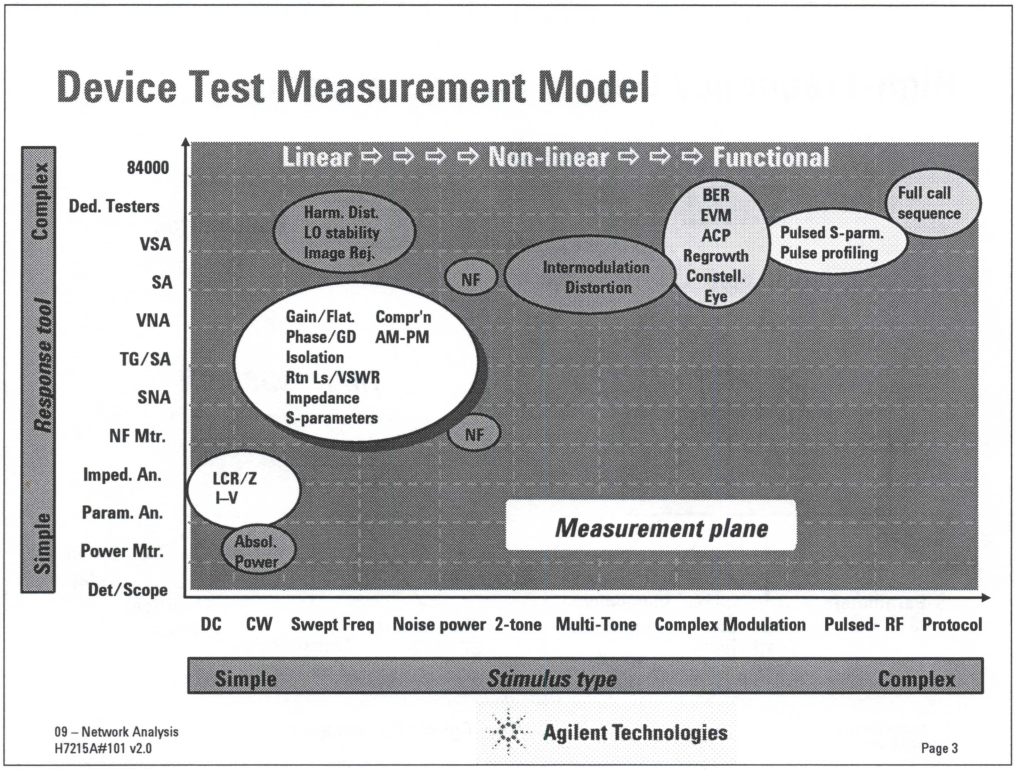


Agilent Technologies



The testing of components, circuits and even systems generally begins with proof of their linear characteristics. The purpose is to confirm the basic, intended operation of the unit. These basic tests might include the measurement of frequency, bandwidth, power, and gain. Once the device has been shown to function as intended, tests of its non-linear characteristics can be performed. Here we are measuring the unwanted qualities of the component. Distortion, frequency response flatness, harmonics are examples of testing for non-linear.

The primary function of a network analyzer is to determine the the linear response of a device under test (DUT) to sinewave stimulus.

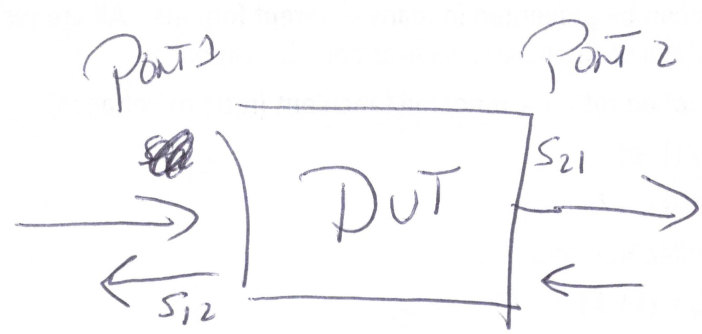


Here is a more complete testing model showing the progression of tests from linear to non-linear and finally, functional testing.

The network analyzer takes its place early in the process employing swept-frequency sinewave stimulus for linear testing and some limited capability for non-linear test.

# of S params = n^2

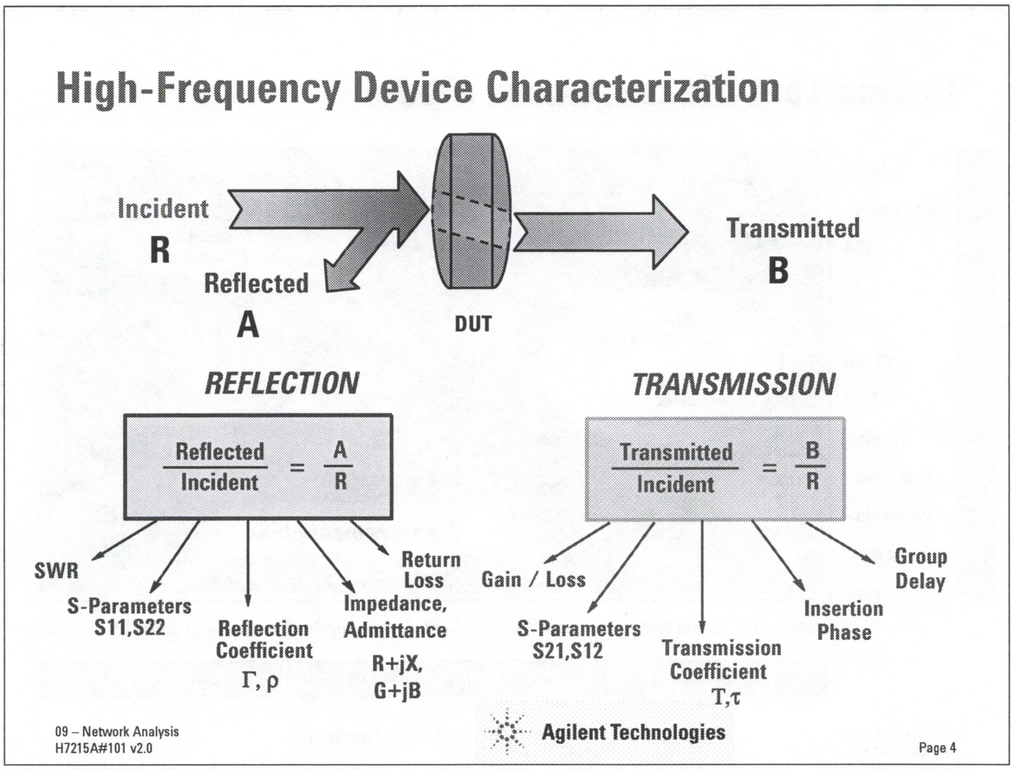
n = number of ports



Meas @ 1
Coming from 2

S_{21}

Meas @ Port 2
Coming in at Port 1



Linear assessment of a high-frequency, RF or microwave device has 2 basic tests:

When a DUT has an RF or microwave sinewave stimulus incident upon it, there are only two ways for the signal to go; 1. Reflected from the device or 2. Admitted into the device. To find out what happened to the admitted wave, we can measure what, if any, signal was transmitted through the device. Both of these measurements are actually ratios of the measured signal relative to the signal incident upon the device.

Both of these ratios can be presented in many different formats. All are just arithmetic manipulations of the basic scalar or complex ratio.

Given the scalar reflection ratio r = reflected/ incident (ratio of voltages)

- SWR= $(1+\rho)/(1-\rho)$
- Return loss(dB) = $-20 \log \rho$

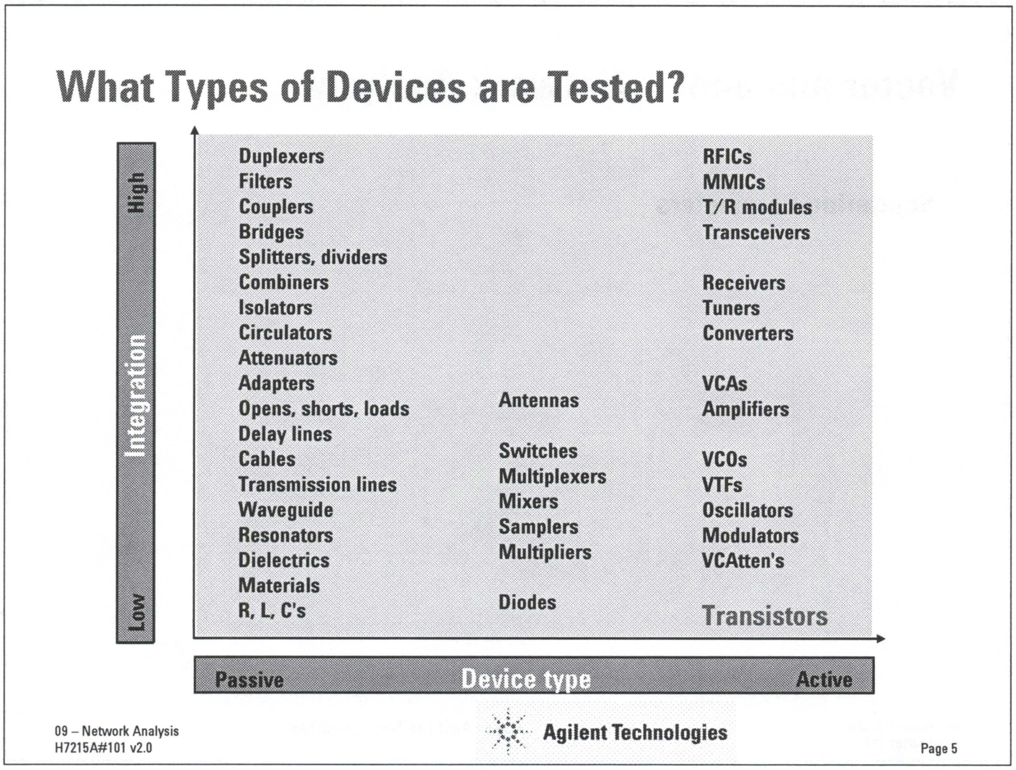
Given the complex reflection ratio Γ:

- Impedance $Z_l = (1+\Gamma)/(1-\Gamma)$
- S11 or S22 = $\Gamma = \rho \,/\underline{\theta}$

For measurements of transmission:

- The complex transmission coefficient = $T = \tau \,/\underline{\theta}$ = transmitted/ incident (ratio of voltages) Gain = $20 \log \tau$; Loss = $-20 \log \tau$ S21 or S12 = T

We will discuss all of these measurement formats at length during this class.



RF and microwave linear testing covers a wide range of items. The most common of these would be components found in most designs of receivers and transmitters like those found in radars and radios. These would include filters, cables and other transmission lines, attenuators, mixers, couplers, splitters and combiners. Linear testing also extends to lower level components such as resistors, capacitors, inductors, and transistors as well as high level components like amplifiers, RFIC's, microwave modules and even whole assemblies like a receiver. A less obvious application would be to test the response of materials to RF and MW energy. Here we can measure dielectric and magnetic characteristics as well as reflectivity, transmission properties, absorption and resonance.

In this class we will test some basic components like filters and transmission lines.

Vector and Scalar Network Analysis

- **Scattering Parameters**

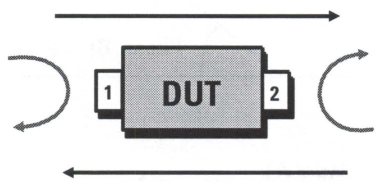
09 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 6

Characterizing a Two-Port Device

- **With signal incident upon either port, what is reflected and transmitted?**



$$\text{Transmission Coefficient} = \frac{\text{Transmitted}}{\text{Incident}}$$

$$\text{Reflection Coefficient} = \frac{\text{Reflected}}{\text{Incident}}$$

09 – Network Analysis
H7215A#101 v2.0
Agilent Technologies
Page 7

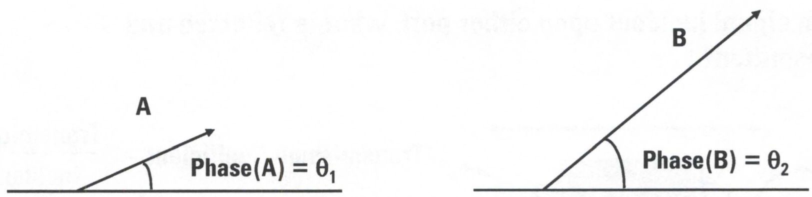
When determining the linear characteristics of a device, what we need to know is fairly simple. Given a sinewave signal incident upon the DUT, what of that signal is reflected from the device or transmitted through it? From the arrows in the illustration, there are four possibilities for this 2-port device. Knowing these four parameters completely characterizes this device for its linear operation.

For an n port device there is a reflection coefficient at each port and a transmission coefficient between each pair of ports in each direction.

$$\text{Transmission Coefficient} = \frac{\text{Transmitted}}{\text{Incident}}$$

$$\text{Reflection Coefficient} = \frac{\text{Reflected}}{\text{Incident}}$$

Signals are Complex Quantities



$$\frac{A}{B} = \frac{Mag(A)\ \angle Phase(A)}{Mag(B)\ \angle Phase(B)} = \frac{Mag(A)}{Mag(B)}\ \angle(\theta_1 - \theta_2)$$

S-Parameter Conventions

- Reflection Measurements = Numbers are the Same - S_{11}, S_{22}
- Transmission Measurements = Numbers are Different - S_{21}, S_{12}

$$\mathbf{S}_{2\ 1}$$

Signal Measured Here (Out of DUT)	Signal Incident Here (In to DUT)

09 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

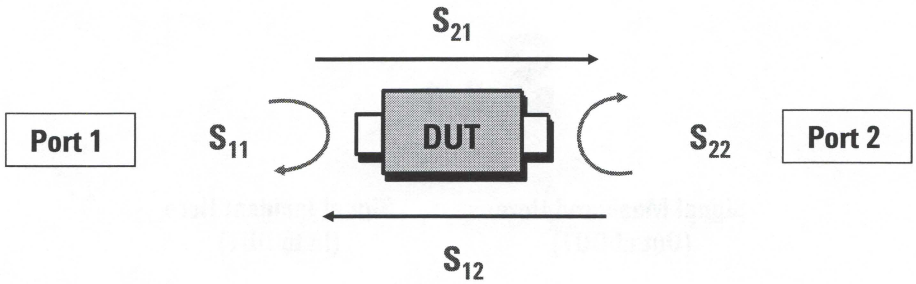
Page 9

To quantify and express these reflected and transmitted coefficients, the scattering or S-parameter were defined. S-parameters give a simple way to quantify and organize these ratios of voltages exiting or scattering from a network relative to the voltages incident upon the network.

S-parameters are always a ratio of two complex (magnitude and phase) quantities. S-parameter notation identifies these quantities using the numbering scheme shown above. The first number refers to the test-device port where the signal is emerging, or another way to look at it, which network analyzer port is the signal being measured. The second number refers to which test-device port the signal is incident or which network analyzer port the signal is coming from. For example, the S-parameter, S11, identifies the measurement as the complex ratio of the signal emerging from port 1 of the device to the signal applied to port 1 of the device (a reflection measurement).

S-Parameters of a Two-Port Device

- **Completely characterize a two port device with four S-parameters.**



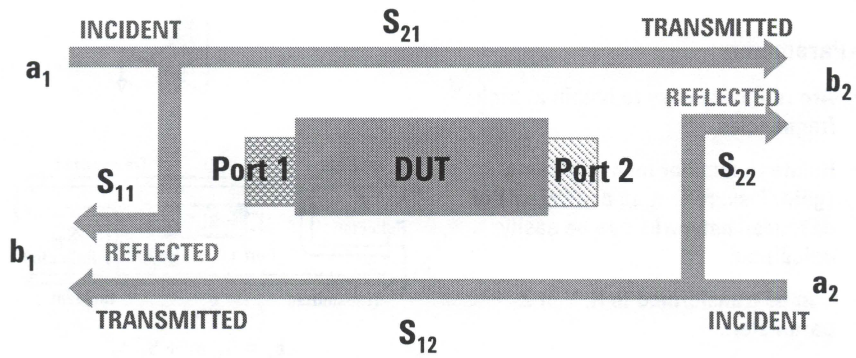
09 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 10

A two-port device or network has four S-parameters. Two of the terms are related to the reflection from the input and output ports of the DUT. The other two terms are related to the transmission through the DUT in the forward and reverse directions. These concepts can be expanded to multi-port devices and the number of S-parameters is a function of 2^n, where n = the number of ports. For example, a four port device would have 16 S-parameters.

S-Parameters



$$b_1 = S_{11}a_1 + S_{12}a_2$$

$$b_2 = S_{21}a_1 + S_{22}a_2$$

09 – Network Analysis
H7215A#101 v2.0

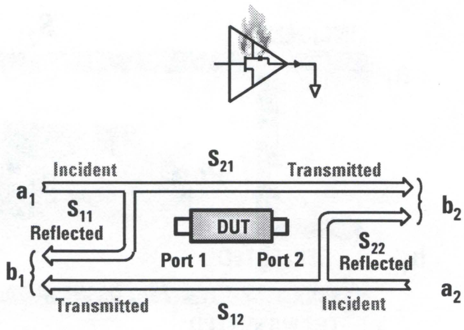
Agilent Technologies

Page 11

For the two port device there are two independent equations may be written, expressing the independent variable b in terms of the dependant variable a. In the above diagram b_1 comprises the sum of a quantity reflected from port 1 and a quantity that is the result of transmission through the device in the reverse direction. The quantities are scaled to be proportional to the voltage wave amplitude and phase such that $|b_n|^2$ = power emerging from the n'th port and $|a_n|^2$ is the power incident on the n'th port.

Why Use S-Parameters?

- S-Parameters:
  - Are relatively easy to obtain at high frequencies
  - Relate to familiar measurements (gain/loss, reflection coefficient) of cascaded networks can be easily calculated.
  - Can be transformed to H, Y or Z parameters.
  - In the form of file data are easily used by CAD programs.



$$b_1 = S_{11} a_1 + S_{12} a_2$$
$$b_2 = S_{21} a_1 + S_{22} a_2$$

09 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 12

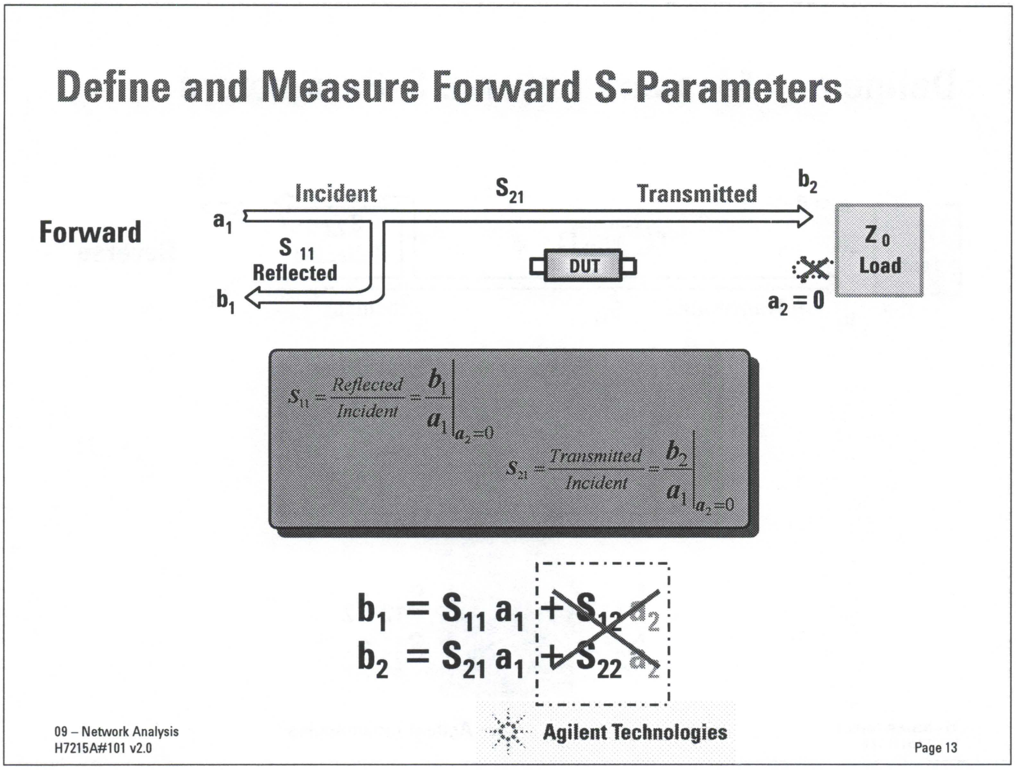
Over the years there have been many methods developed to model the behavior of a circuit or network. H, Z and Y parameters as well as Spice models all have benefits to visualizing and predicting circuit behavior. However, as the circuit frequency climbs higher, many of these models become complex and unruly. The result it is hard to model the network in a straightforward manner.

Scattering (S)-parameters give a simple way to quantify the ratio of voltages exiting or scattering from a network relative to the voltages incident upon the network. By cascading* the parameters, models for individual circuit blocks may be put together to predict the operation of the whole system.

As we will find out these ratios are fairly easy to measure at high frequencies. Difficult measurements of complex short-circuit currents and open-circuit voltages are not required to measure S-parameters.

Since S-parameters are defined as a voltage ratio, their result maps well into familiar RF and microwave characteristics of gain, loss, an reflection coefficient. They are also easily converted to other parameter forms such as H, Y or Z, and this transportability makes S-parameters a good choice for importation into most modern simulation and design software programs.

For students acquainted with matrix algebra cascading involves matrix multiplication of "T" matrices which are derived by a transformation of the s-parameter matrices. See AN#154 "S-Parameter Design."



To define S-parameters, it is assumed that a sinewave is incident upon one port of the device while the other port is terminated by a Z_0 load.

$$s_{21} = \text{Transmission Coefficient} = \frac{\text{Transmitted}}{\text{Incident}} = \left.\frac{b_2}{a_1}\right|_{a_2=0}$$

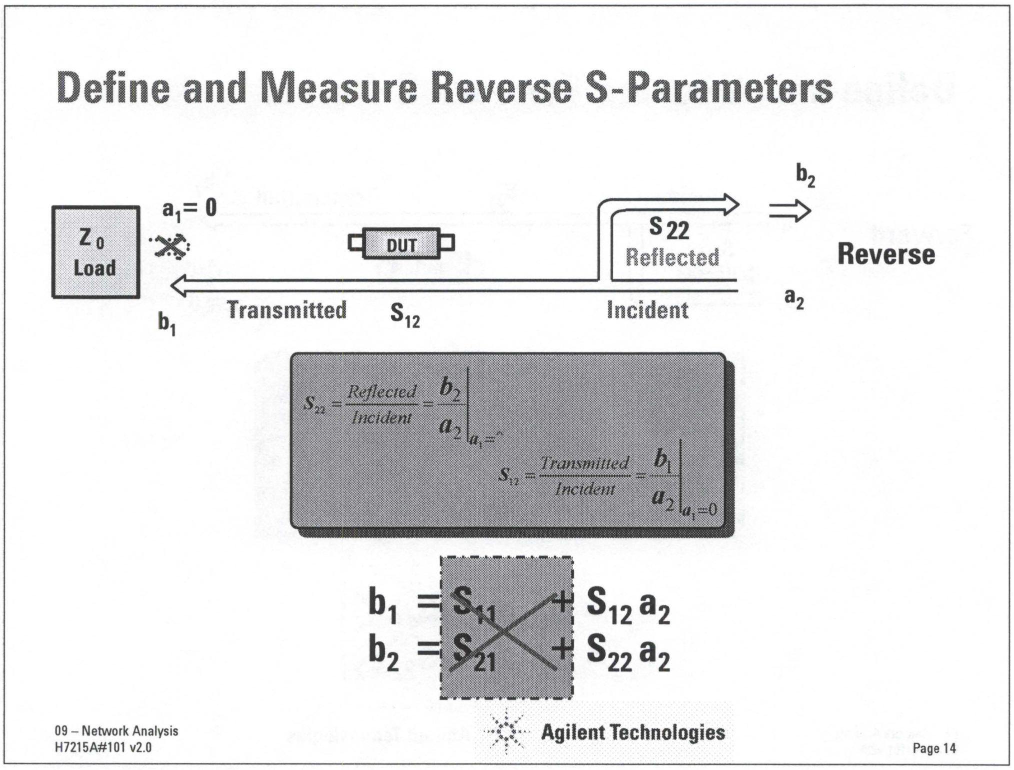
$$s_{11} = \text{Reflection Coefficient} = \frac{\text{Reflected}}{\text{Incident}} = \left.\frac{b_1}{a_1}\right|_{a_2=0}$$

For a 2-port device, there will be 2 scenarios:

- The stimulus is incident upon port 1 of the device. This will be call the *forward* condition, defined above.
- The stimulus is incident upon port 2 *(reverse).*

The test system must be able to sample the incident signal seen as quantity a_1 in the forward direction and a_2 in the reverse direction. The setup must also measure the signal reflected and transmitted (b). Lastly, to yield S-parameters, the system must be able to take the complex (magnitude and phase) ratio of the reflected or transmitted signal versus the incident signal (b/a). The measurements and ratios must be complex to produce true S-parameters.

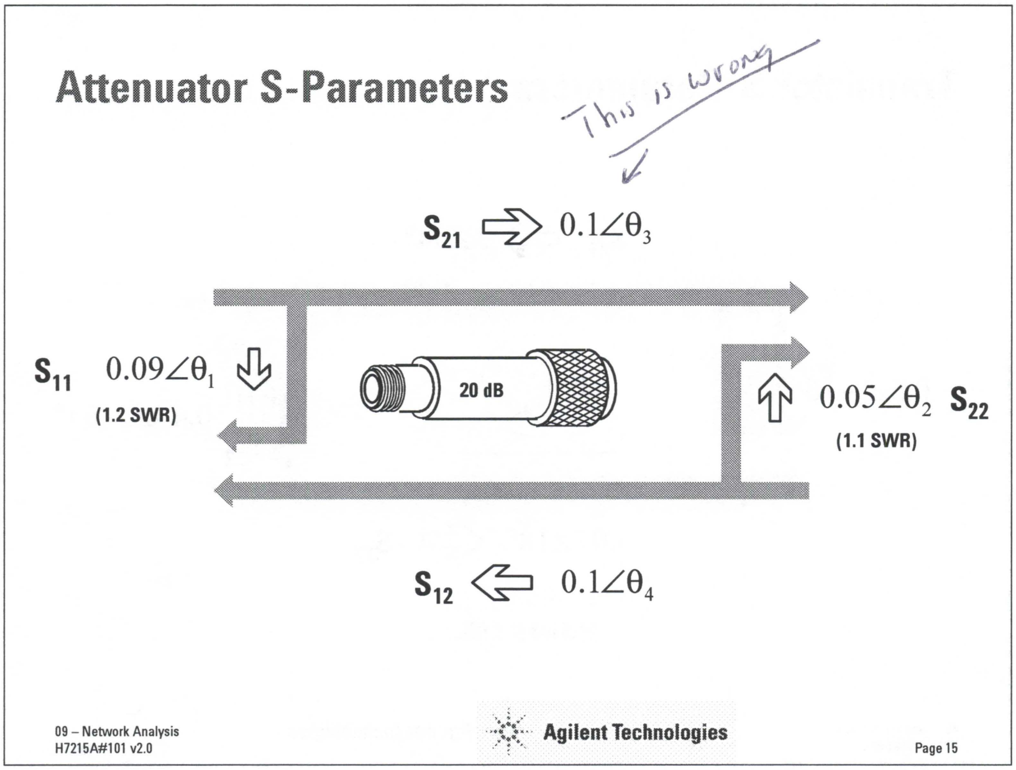
In this course, we will learn how the test system performs this task.



With the stimulus at port 2, (reverse condition):

$$s_{12} = \text{Transmission Coefficient} = \frac{\text{Transmitted}}{\text{Incident}} = \frac{b_1}{a_2}\Big|_{a_1=0}$$

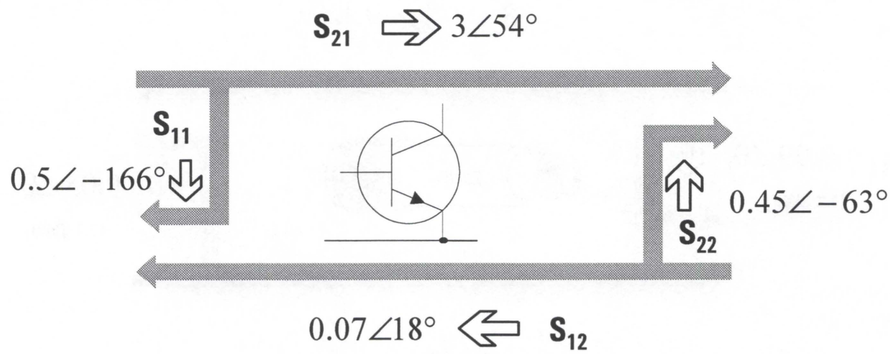
$$s_{22} = \text{Reflection Coefficient} = \frac{\text{Reflected}}{\text{Incident}} = \frac{b_2}{a_2}\Big|_{a_1=0}$$



Let's look at some examples of S-parameters for typical devices. Shown here is a 20 dB attenuator. Remembering that S-parameters are voltage ratios, it makes sense that the forward transmission S-parameter, S_{21}, would be .1 with some phase shift from incident to output. An attenuator is just a voltage divider, so its reverse transmission characteristics, $S_{12,}$ will be similar (a bi-lateral device). One of the attenuator's design goals is to be well matched to Z_0. Therefore, the reflection S-parameters, $S_{11,}$ and S_{22} will be small (0.09 and 0.05 respectively).

Once these linear ratios have been measured, they can easily be converted to other common formats such as standing wave ratio (SWR), or expressed as gain or loss parameters in dB.

Transistor S-Parameters



2N5103 @ 1 GHz

09 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

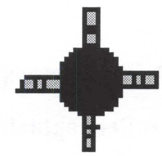
Page 16

A transistor will have a very different set of S-parameters from an attenuator. Notice, in the forward direction, S_{21} is a positive whole number denoting device gain, where in the reverse direction the device is lossy. The reflection S-parameters indicate a large percent of the incident signal is reflected. This is typical of transistors to have a poor match and require a matching networks to couple signals into and out of the device.

We can start to see the value of making these measurements. From here the designer can create matching networks for the transistor to maximize signal transfer, gain and stability of the combined or cascaded networks.

S-Parameter Transistor Data

AT- 41486



!AT-41486!
S AND NOISE PARAMETERS at Vce=8V Ic=10mA. LAST UPDATED 07-21-92
ghz s ma r 50

F GHz	S11		S21		S12		S22	
	Mag.	Angle	Mag	Angle	Mag	Angle	Mag	Angle
0.1	.74	-38	25.46	157	.011	68	.94	-12
0.5	.59	-127	12.63	107	.031	47	.60	-29
1.0	.56	-168	6.92	84	.041	46	.49	-29
2.0	.62	152	3.61	56	.058	43	.42	-39
3.0	.64	130	2.41	37	.078	52	.39	-50
4.0	.71	113	1.80	16	.106	48	.35	-70
5.0	.77	99	1.42	-4	.139	43	.35	-98
6.0	.81	87	1.13	-22	.170	34	.35	-131

S-Parameter Summary

- **S-Parameters are:**
  - **Conceptually Simple**
  - **Precise and Comprehensive**
  - **Easy to Measure and Interpret**
  - **Analytically Convenient**
    - **For importation to CAE programs**
    - **For Flow Graph Analysis**

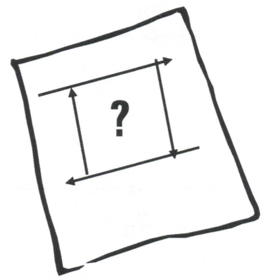
09 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 18

Too summarize, S-parameters provide a simple way to fully characterize the linear behavior of a network. They are straightforward to measure and understand. Most importantly they lend themselves well to conversion to other formats, they cascade well for large-system simulation, and import readily to design and simulation programs.

What are Flow Graphs?



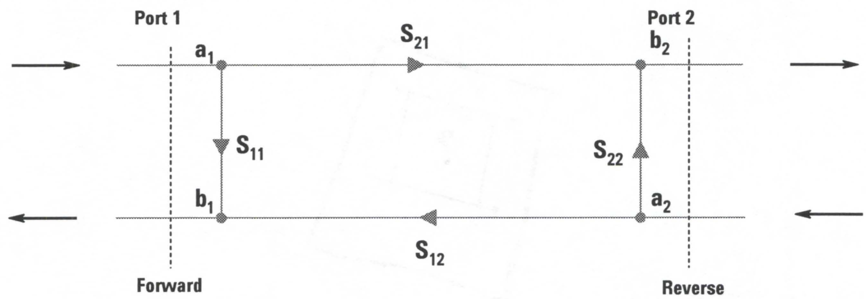
09 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 19

Flow Graph Representation

S-Parameters



$$\begin{bmatrix}\text{Number of} \\ \text{S Parameters}\end{bmatrix} = \begin{bmatrix}\text{Number of} \\ \text{Ports}\end{bmatrix}^2$$

09 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 20

Just as S-parameters are key tools for RF and microwave design and analysis, flow graphs also play a key role in developing and understanding system behavior.

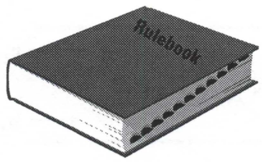
The illustration shows a flow graph for a 2-port network. Each port has 2 nodes. One node represents the signal entering the signal entering the port and the other node represents the signal exiting the port. Lines that connect nodes are branches. Each branch has an arrow and a value corresponding to an s-parameter. Signals will only flow in the direction of the arrows. The number of branches in the flow graph equals the number of the network's s-parameters, which is equal to the square of the number of network ports.

NOTE: In S-parameter theory the wave incident on the network is the "a" wave and the wave exiting from a network is the "b" wave, when networks are cascaded the exiting wave from one network will be the incident wave on the next.

The units of the waves are proportional to voltage such that $|a|^2 =$ power incident and $|b|^2 =$ power emerging. (Power to a reference load Z_0)

Using a Flow Graph

- **Rulebook**
  - **Loop rules**
  - **Paths**
  - **Etc.**

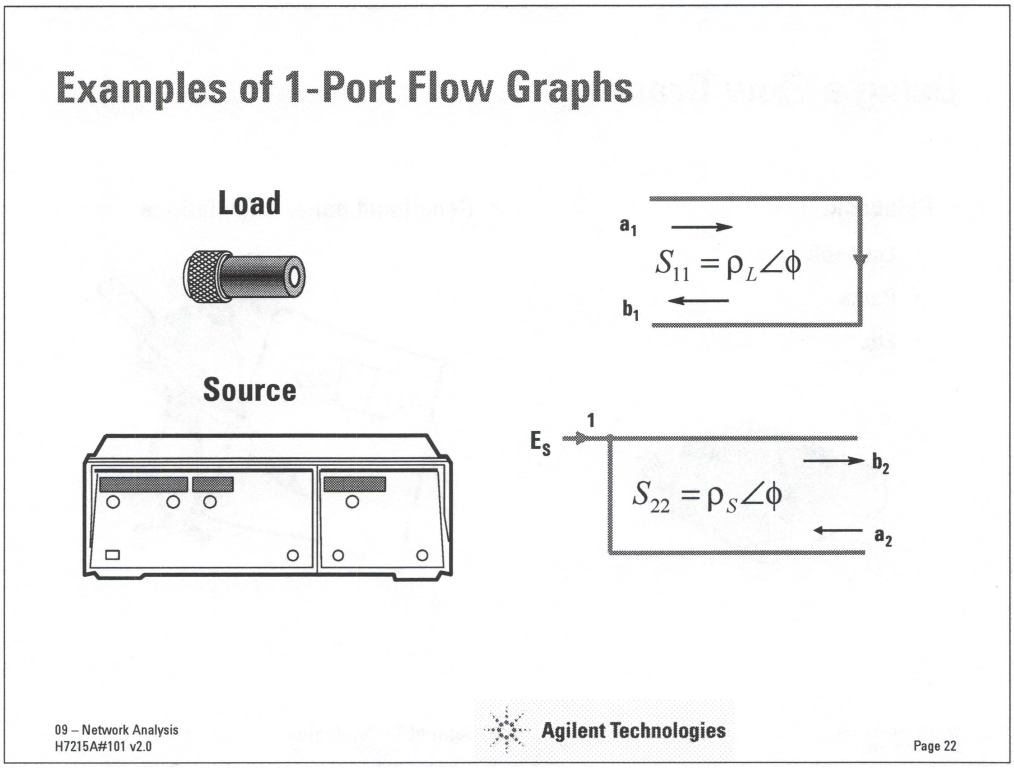


- **Pencil and paper calculations**

09 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

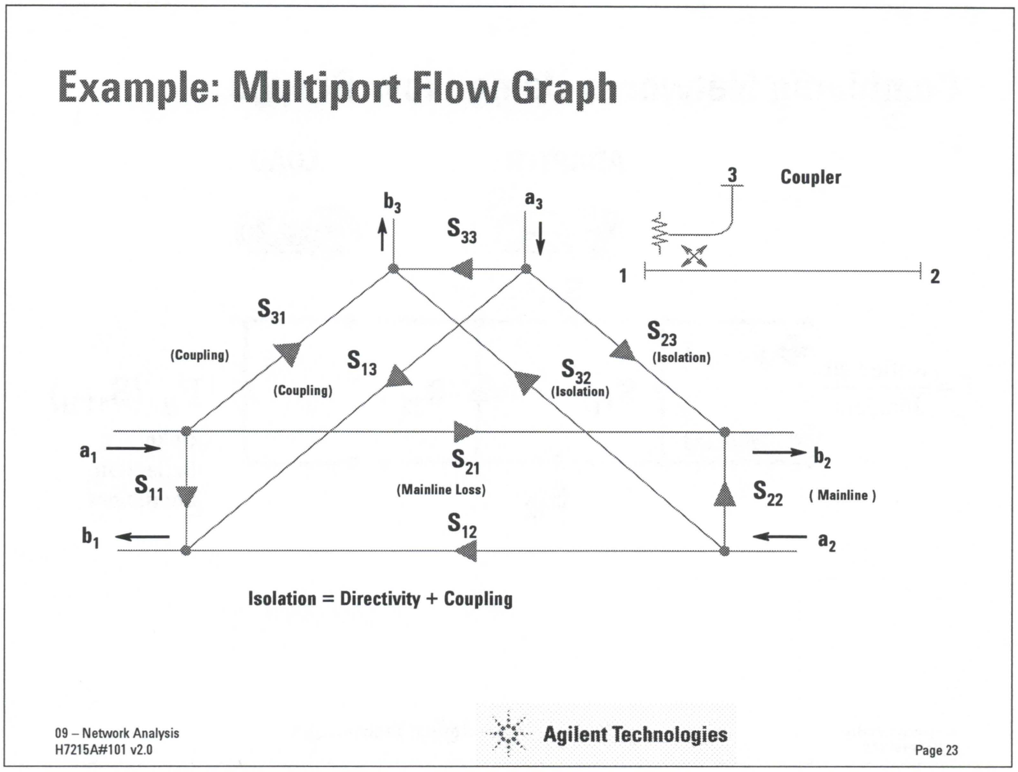
Page 21



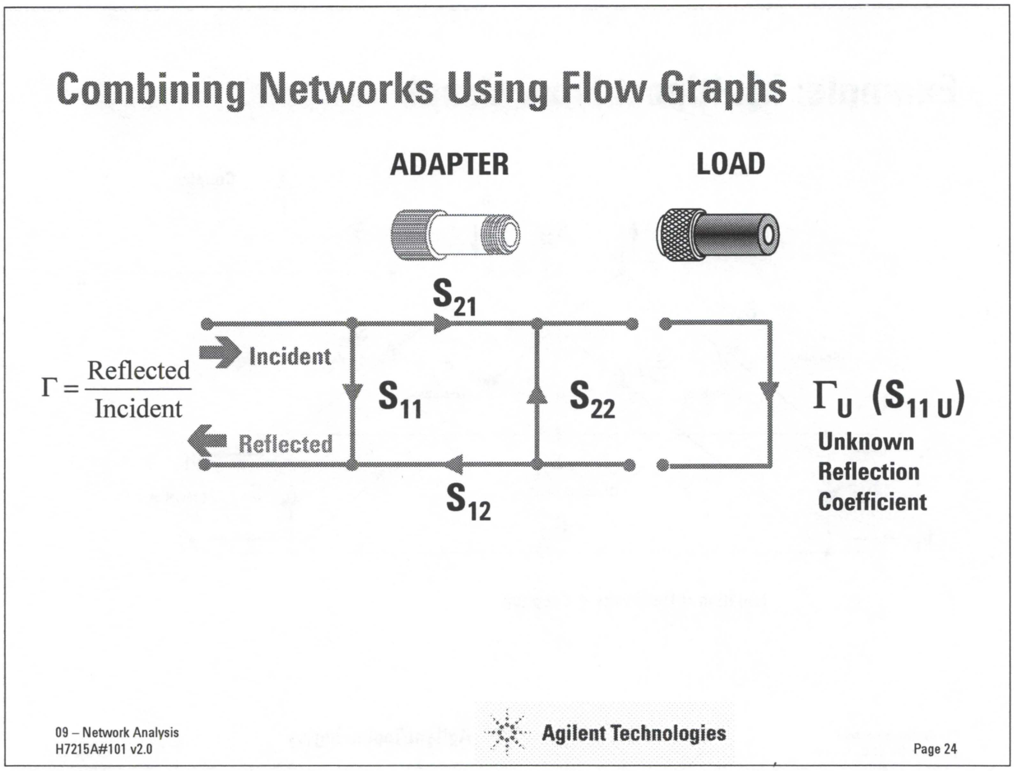
A one-port network has one s-parameter. Here are two example 1-port networks, a load and a source, and their corresponding flow graphs. Each flow graph shows how the incident signal will be modified by the network and exit the network as a reflected signal.

For the load $b_1 = S_{11}a_1$

For the source $b_1 = E + S_{22}a_1$



A multi-port network has N^2 number of s-parameters where N is the number of ports. Here is shown the flow graph for a 3-port coupler and its nine s-parameters. The names that accompany the s-parameters are associated with the function of the device.

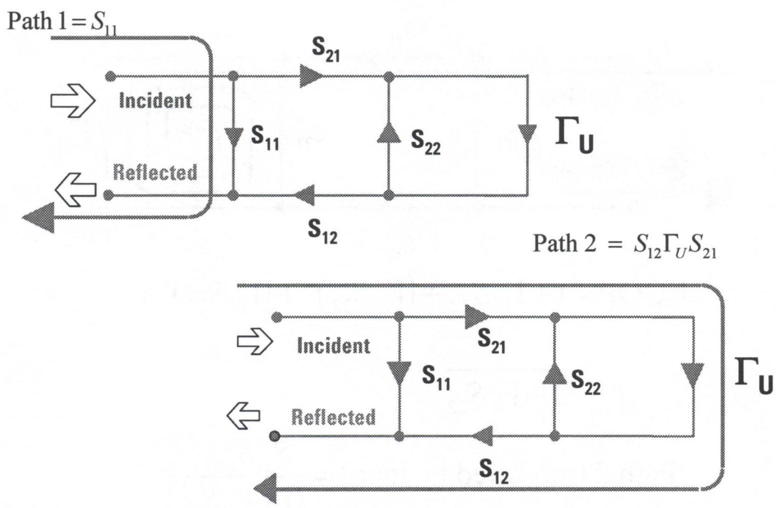


Because flow graphs are used so frequently in RF and microwave documents, it is important to appreciate the nature of the representation and technique. We shall derive the effect of an adapter by using the flow graph rules, this simple but very important result sets the stage for understanding error correction which comes later in this section.

Flow graph techniques can be used to determine the overall response of a network once the component parts are known. As an example, let's use an adapter to convert a Z_0 termination to another connector type. We know intuitively that the load is not improved by putting something in front of it. Using the rules* of flow graph analysis, the composite response of the combined networks can be found.

** The rules may be found in AN#154 "S-Parameter Design."*

Flow Graph Paths



09 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

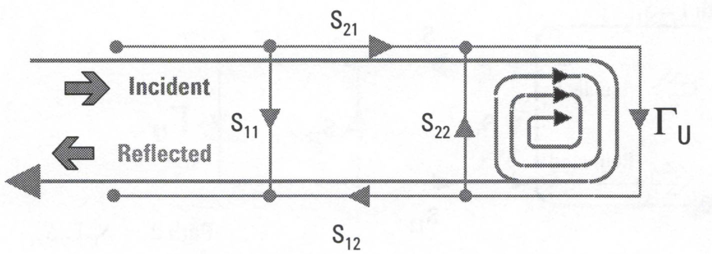
Page 25

Flow graph analysis is performed using a book of rules. The first task is to establish the independent paths through the graph.

Energy flows from the source to sink in the direction of the arrowheads. Thus, one obvious path for energy incident at port one is the direct path through the adapter's S_{11}

A second path or direction energy may flow in is through S_{21}, then through the Γ of the load, and then back out through the S_{12} of the adapter. The combined value of this path is the product of each branch or ($S_{21} \times \Gamma_u \times S_{12}$). Notice that while energy cannot flow out of S_{12} into S_{22}, nothing prevents it from coming up S_{22} and back into Γ_u.

Flow Graph Paths (continued) Apply the Rules!



$$LOOP = 1 + \Gamma_U S_{22} + (\Gamma_U S_{22})^2 + (\Gamma_U S_{22})^3 + \ldots$$

$$= \frac{1}{1 - \Gamma_U S_{22}}$$

$$\text{Path 2 (modified by loop)} = \frac{S_{21} S_{12} \Gamma_U}{(1 - \Gamma_U S_{22})}$$

09 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 26

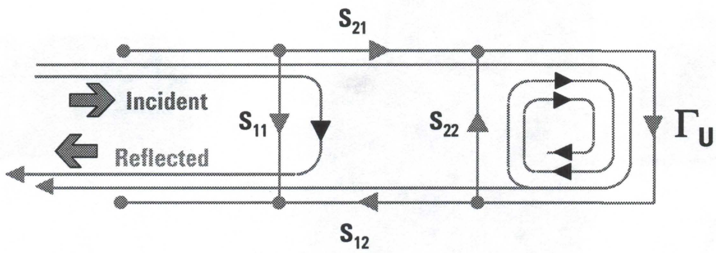
A flow graph reduction technique known as "Mason's Nontouching-loop Rule" says that the overall response of the network is the sum of the paths which energy can flow through in the network.

Some of the energy does flow in that direction and re-circulates in a loop. Each time it transverses the loop a product of the coefficients is generated (in other words, an in finite series). A mathematical identity converts the infinite series to a simple expression which is multiplied by the S_{21}, S_{12} and Γ_u branches to give an overall value for path two.

Algebra Refresher:

$$\text{if } Sum = 1 + x + x^2 + x^3 + \ldots$$
$$= 1 + x(1 + x + x^2 + x^3 + \ldots)$$
$$= 1 + x.Sum$$
$$\therefore Sum(1 - x) = 1$$
$$Sum = \frac{1}{1 - x}$$

Composite Response



$$\frac{\text{Reflected}}{\text{Incident}} = \text{Path 1} + \text{Path 2 (modified by loop)}$$

$$= S_{11} + \frac{S_{21}S_{12}\Gamma_U}{(1-\Gamma_U S_{22})}$$

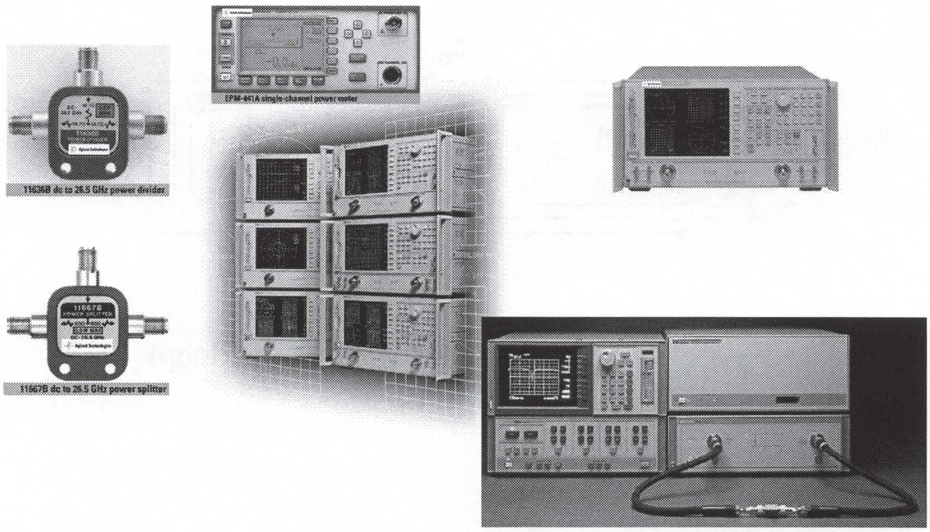
09 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 27

Now for the sum. This simple mathematical relationship is an exact solution to any two-port network terminated in one port. In our example, it is the reflection coefficient of the adapter/load combination. Notice the prominence of the S_{11} term. No matter how low the load reflection coefficient is, the adapter's input reflection S_{11} will seriously degrade the "new" load even when it is comparable in value to Γ_υ. Note that all four s-parameters are required to calculate just one. The importance of this statement will be seen in later topics.

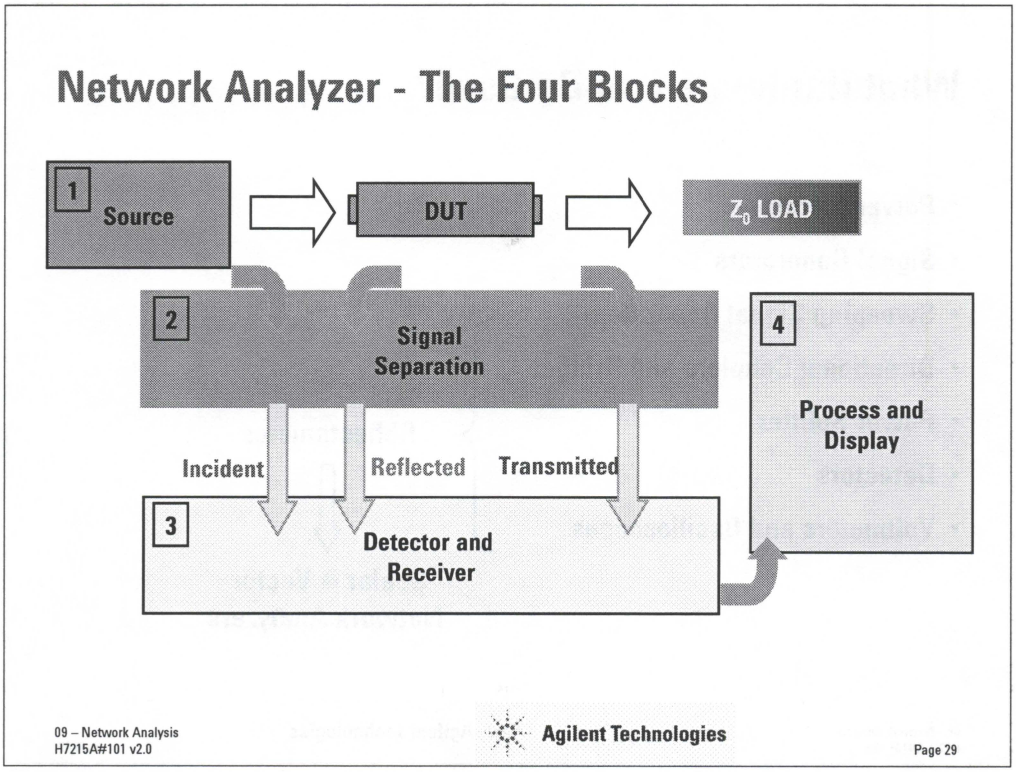
Hardware



09 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 28


Network Analyzer - The Four Blocks
1
Source
DUT
Z_0 LOAD
2
Signal
Separation
4
Process and
Display
Incident
Reflected
Transmitted
3
Detector and
Receiver
09 – Network Analysis
H7215A#101 v2.0
Agilent Technologies
Page 29

What Hardware Can Be Used?

- Power Meters
- Signal Generators
- Sweeping Signal Generators
- Directional Couplers and Bridges
- Power Splitter
- Detectors
- Voltmeters and Oscilloscopes

Reflectometer

Scalar & Vector Network Analyzers

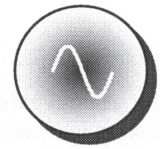
09 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

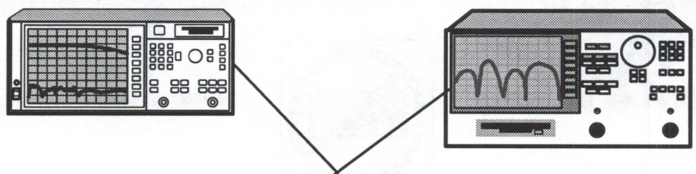
Page 30

To measure signals reflected from and transmitted through a network, a host of different tools and methods may be used. One approach is to make absolute power measurements and compute the ratios later. Another is to assemble some common components into a system known as a reflectometer. For greater convenience, speed and accuracy, a dedicated reflectometer can be acquired. The ultimate in accuracy and analysis capability is the vector network analyzer.

Source



- **Supplies stimulus for system**
- **Swept frequency or power**
- **Traditionally NAs used separate source**
  - **Open-loop VCOs**
  - **Synthesized sweepers**
- **Most analyzers sold today have integrated, synthesized sources**
- **(The internal source may not be suitable as a general purpose source)**

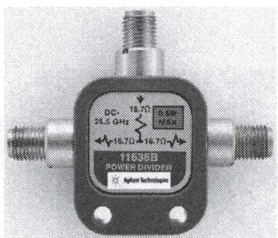


1

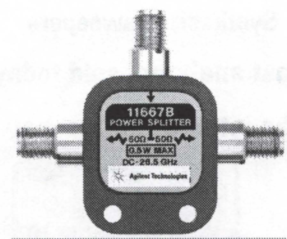
Integrated, synthesized sources

Dividers and Splitters

- **Power Divider**
  - **Use for general purpose signal monitoring independent devices**
- **Power Splitter**
  - **Use for NWA ratio or**
  - **Automatic Level Control (ALC)**



11636B dc to 26.5 GHz power divider



11667B dc to 26.5 GHz power splitter

09 – Network Analysis
H7215A#101 v2.0

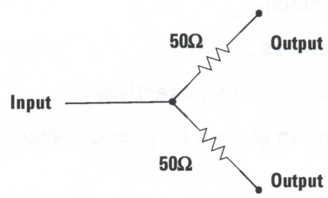
Agilent Technologies

Page 32

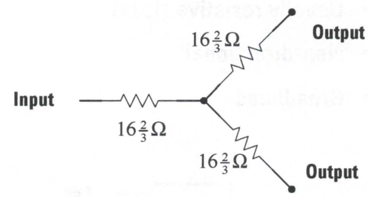
They look the same, but the splitter or divider must be chosen for the proper applications. NEVER use the three-resistor power divider as a network analysis ratio reference.

Splitters and Dividers

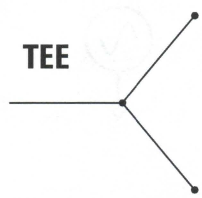
- Splitter



- Divider



- Leveling
- Ratioing



- Simple Splitting (2 Independent Devices)
- At high frequencies poor VSWR, Resonant
- Not Recommended!

09 – Network Analysis
H7215A#101 v2.0

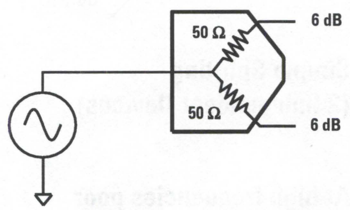
Agilent Technologies

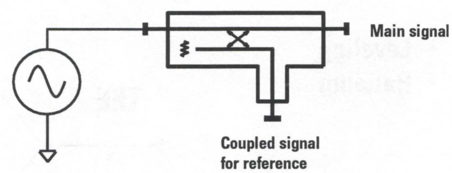
Page 33

NA Hardware - Signal Separation

Measures incident signal for reference

- **Splitter**
  - **Usually resistive (loss)**
  - **Non-directional**
  - **Broadband**
- **Coupler**
  - **Directional**
  - **Low loss**
  - **Good isolation, directivity**
  - **Hard to get low freq performance**





2

09 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 34

In this section, we will cover the signal separation block. The hardware used for this function is generally called the "test set". The test set can be a separate box or integrated within the network analyzer.

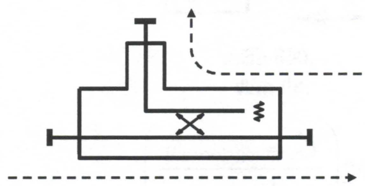
There are two functions that our signal-separation hardware must provide. The first is to measure a portion of the incident signal to provide a reference for ratioing. This can be done with splitters or directional couplers. Splitters are usually resistive. They are non-directional devices (more on directionality later) and can be very broadband. The trade-off is that they usually have 6 dB or more of loss in each arm.

Directional couplers can be built with very low loss (through the main arm) and good isolation and directivity. However, it is hard to make them operate at low frequencies. This can be a problem in RF network analyzers, where low frequency coverage is important.

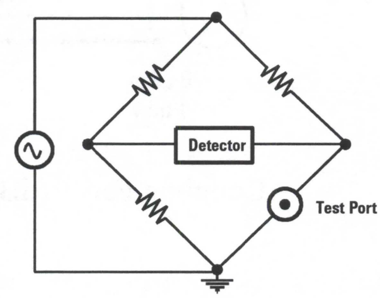
Signal Separation

Separating incident and reflected signals

- **Coupler**
  - **Directional**
  - **Low loss**
  - **Good isolation, directivity**
  - **Hard to get low freq performance**



- **Bridge**
  - **Used to measure reflected signals only**
  - **Broadband**
  - **Higher loss**



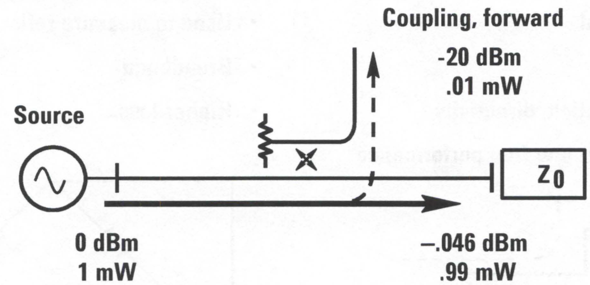
2

The second function of the signal-splitting hardware is to separate the incident (forward) and reflected (reverse) traveling waves at the input of our DUT. Again, couplers are ideal in that they are directional, have low loss, and high reverse isolation. However, due to the difficulty of making truly broadband couplers, bridges are often used.

Bridges are very useful for measuring reflection because they can be made to work over a very wide range of frequencies. Because they are resistive devices, they exhibit an excellent impedance match which is important to minimize measurement uncertainty. The main trade-off is that they exhibit more loss to the transmitted signal, resulting in less power delivered to the DUT for a given source power.

Forward Coupling Factor

- **Example of 20 dB Coupler:**



$$\text{Coupling Factor(dB)} = -10\log\left(\frac{P_{\text{coupling forward}}}{P_{\text{incident}}}\right)$$

2

09 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

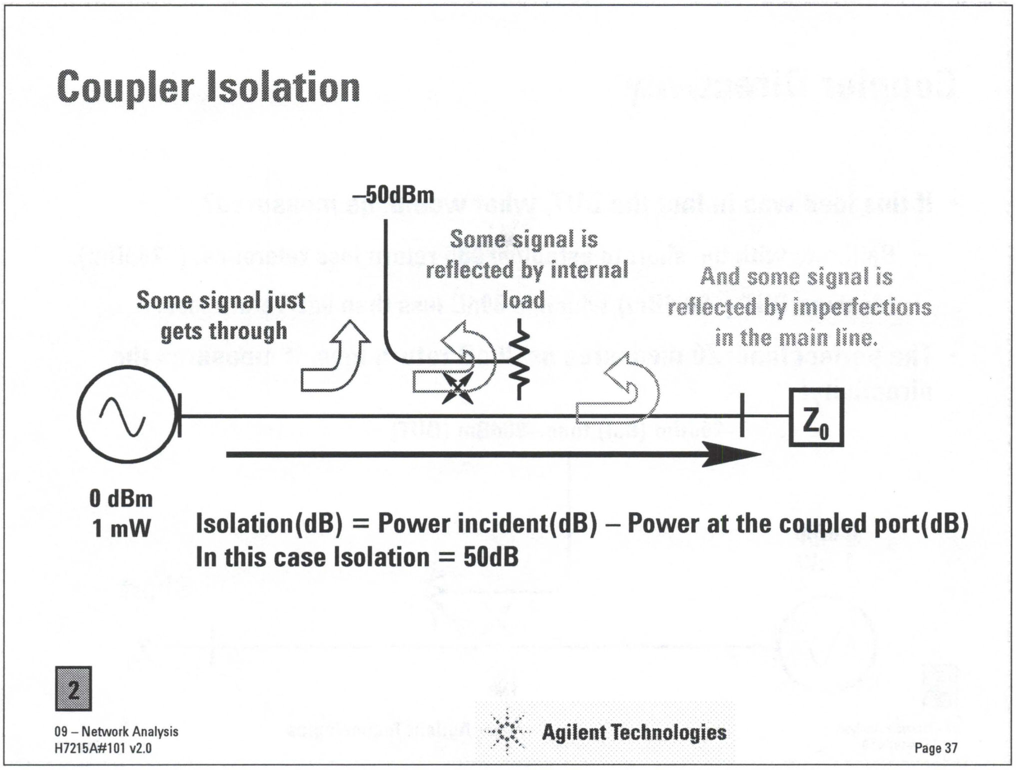
Page 36

The directional coupler measures (couples) a portion of the signal traveling in one direction only. The signal flowing through the main arm is shown as a solid line, and the coupled signal is shown as a dotted line.

The signal appearing at the coupled port is reduced by an amount known as the coupling factor. This is measured by placing the coupler in the forward direction and measuring the power at the coupled port, relative to the incident power:

$$\text{Coupling Factor (dB)} = -10 \log (P_{\text{coupling forward}} - P_{\text{incident}})$$

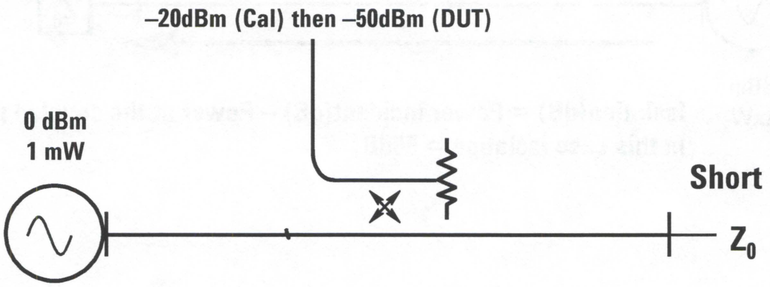
In this example of a 20 dB directional coupler, the level of a signal at the coupled port is 20 dB below that of the input port. The loss through the main arm is only .046 dB. There are also frequency response terms associated with the main-arm response and the coupling factor, expressed in terms of $\pm$ dB.



A perfect coupler would have infinite isolation, that is in the reverse direction, no signal would appear at the coupled port.

Coupler Directivity

- **If this load was in fact the DUT, what would be measured?**
  - **Calibrate with the short to establish 0db return loss reference. (–20dBm)**
  - **Measure DUT (–50 dBm) which is 30dB less than 0dB return loss.**
- **The perfect load; Z0 measures as 30dB return loss, it measures the directivity!**



2

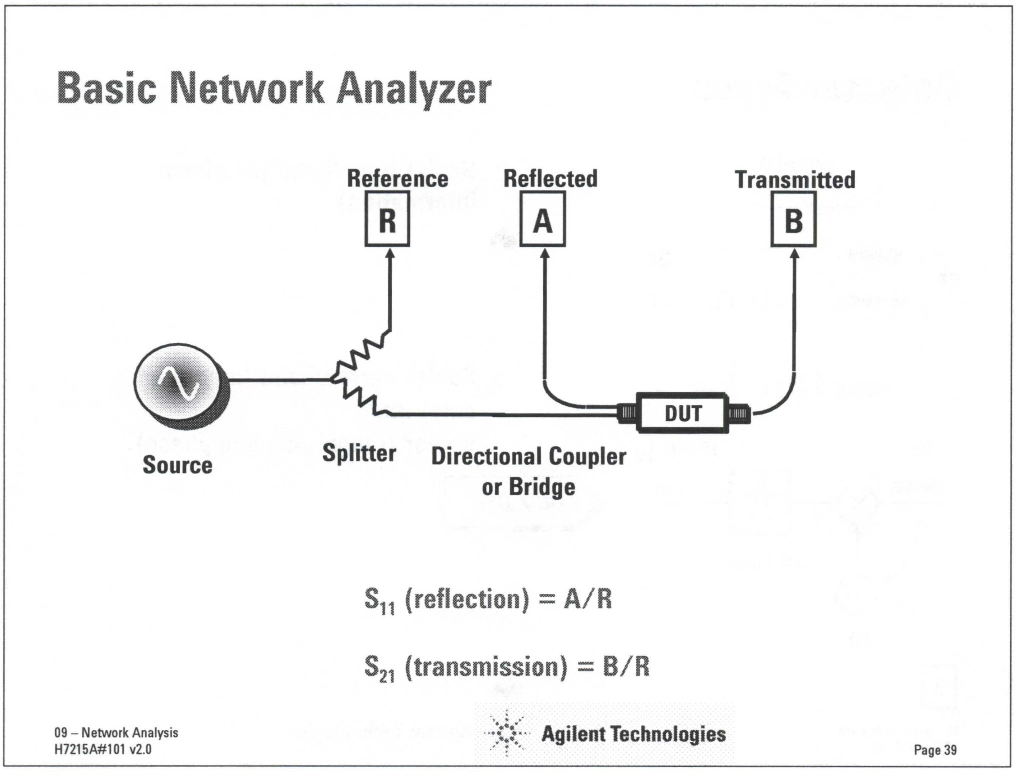
Coupler Directivity

This parameter is a combination of coupling and isolation.

For a coupler that is to be used for reflection measurements the directivity is a most important parameter. The cost of a coupler will be a function of it's directivity.

- Directivity(dB) = Isolation(dB) – Coupling(dB)
- In this case Directivity = (50 – 20) = 30dB

From this example we can deduce that the directivity sets a limit on the ability of a coupler to measure reflections. It is therefore the limit of a scalar network analyzer's ability to measure return loss. We shall see that this limitation may be overcome be vector error correction, however, it is still an important parameter for the couplers used for reflection measurement in a VNA.

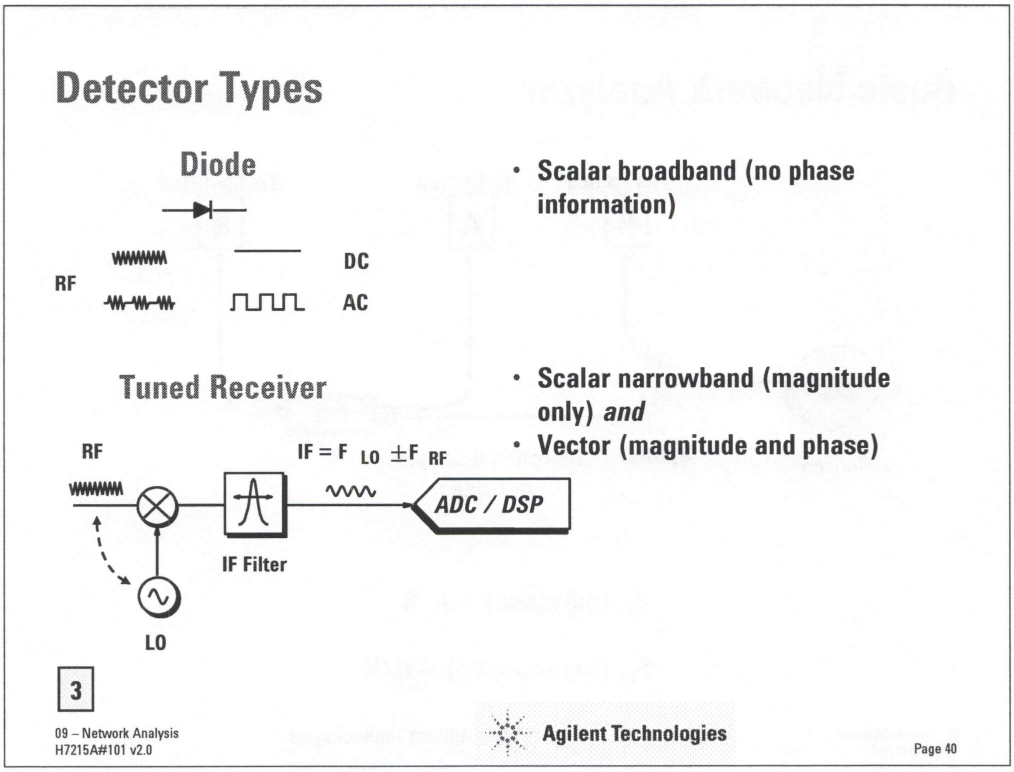


Now let's put some components together to build a reflectometer. A reflectometer will manage the RF and microwave signals for the measurement of reflection and transmission of a DUT. Let's see how it works.

Starting with a source, usually a RF or microwave signal generator or sweep generator, the signal is divided into two equal signals by a splitter. One of these signals called the "reference" and is shown here as "R". This represents the signal incident upon the DUT. The other signal is sent on to the DUT and passes through a directional coupler or bridge on its way the the DUT. Signals reflecting off of the DUT re-enter the coupler and (subject to the coupling factor) are measured at the coupled arm "A." Signals transmitted through the DUT are shown here as quantity "B."

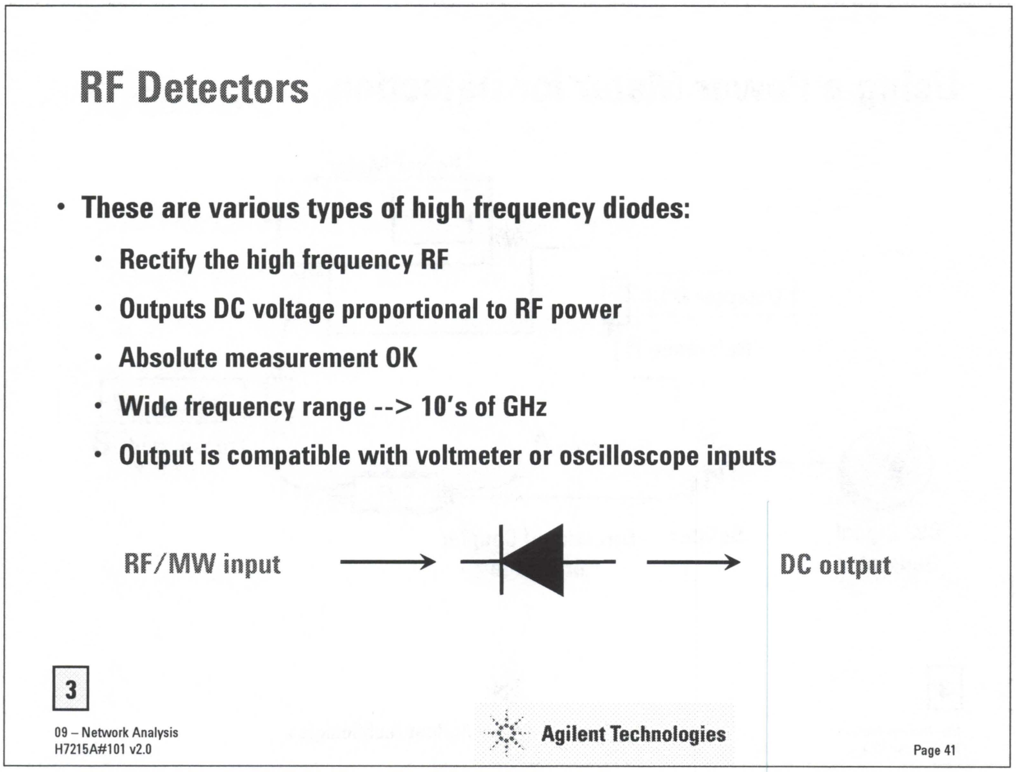
- The reflection ratio or coefficient is simply A/R.
- The transmission coefficient is B/R.

These things apply to both scalar analysis and vector analysis. The difference is how the measurements of "R", "A", and "B" are made. If simple detection is used then only scalar magnitudes are available for the ratioing process. However, if a vector receiver is available to measure these quantities, it is then possible to compute the complex (vector) coefficients for reflection and transmission. This simple diagram is the basis for any network analyzer system. Let's look at different schemes to implement the reflectometer.

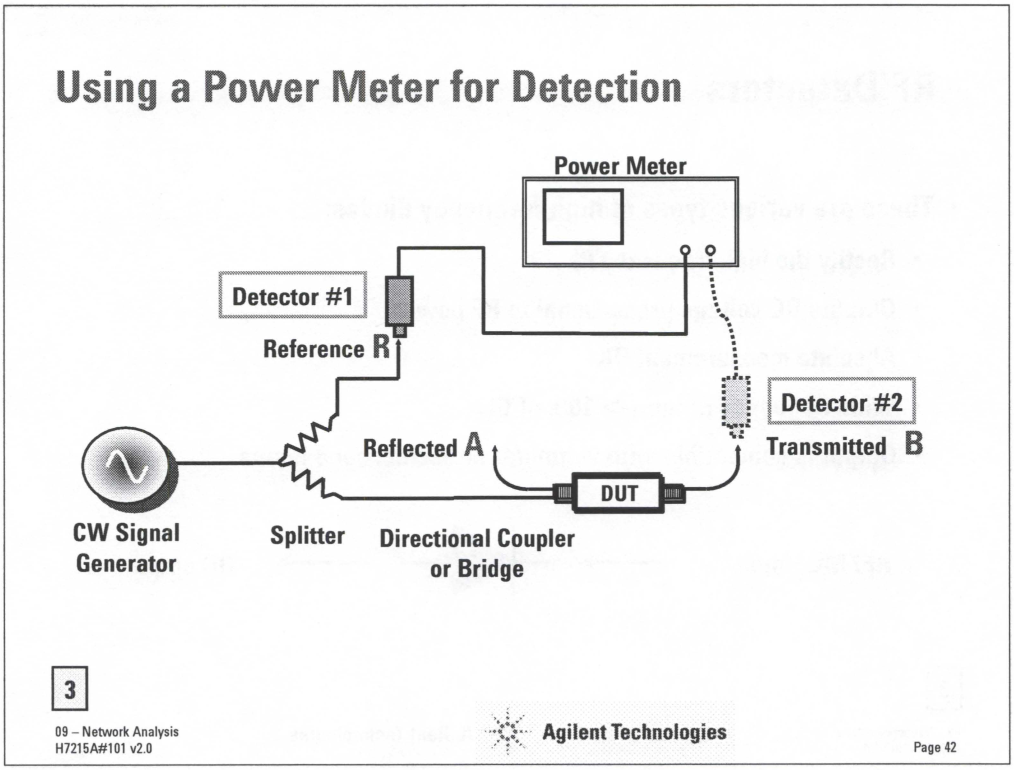


There are two basic ways of providing signal detection in network analyzers. Diode detectors convert the RF signal level to a proportional DC level. If the signal is amplitude modulated (AC detection), the diode strips the RF carrier from the modulation. Diode detection is inherently scalar, as phase information of the RF carrier is lost.

The tuned receiver translates the RF to a lower "intermediate" frequency (IF). The LO is either locked to the RF or the IF signal so that the receivers in the network analyzer are always tuned to the RF signal present at the input. The IF signal is bandpass filtered, which narrows the receiver bandwidth and greatly improves sensitivity and dynamic range. Modern analyzers use an analog-to-digital converter (ADC) and digital-signal processing (DSP) to extract magnitude and phase information from the IF signal. The tuned-receiver approach can be used in scalar or vector network analyzers.

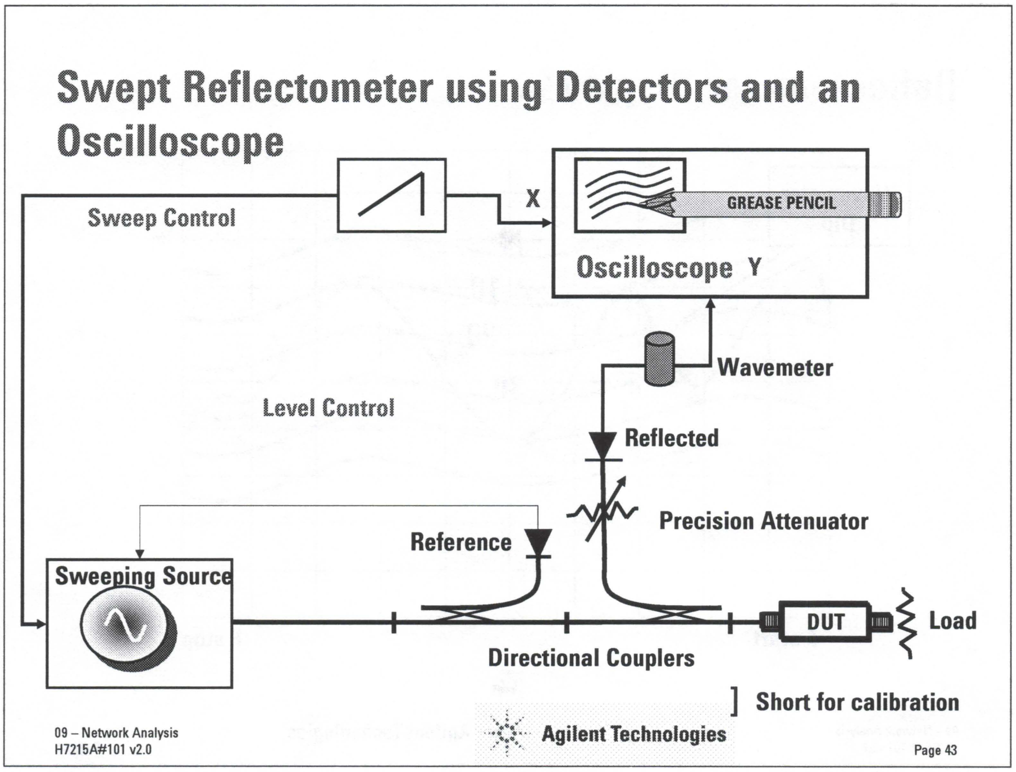


The RF detector is a tool very basic to all applications of RF and microwave work. Detectors can employ a number of different technologies in germanium, silicon, and gallium arsenide, but all of these function as a diode to rectify the high-frequency RF and yield a DC output that is proportional to the power of the RF input. Detectors offer a low-cost, and broadband way to sense RF power. The DC output can then be indicated on low cost instruments like voltmeters or oscilloscopes. Some vector analyzers have diode detectors in addition to the tuned receiver to facilitate absolute power measurement.



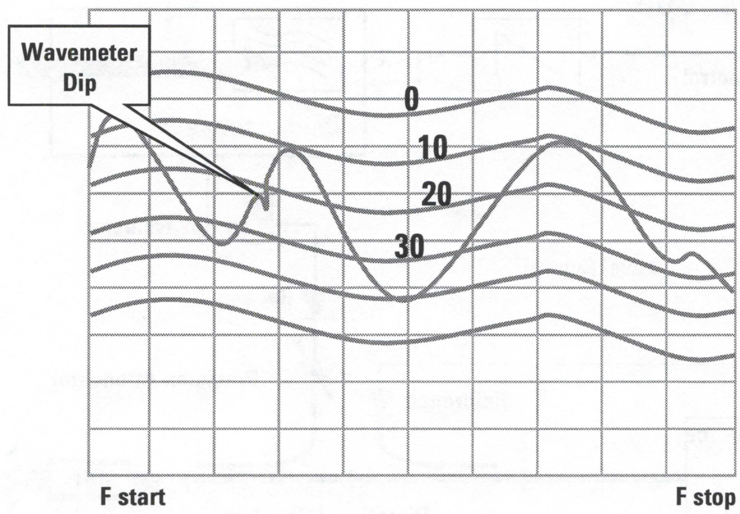
One way to measure the output ports of the reflectometer would be to use the sensors of a power meter. A single sensor could be moved or multiplexed to the three measurement ports and the measurement ratios for transmission and reflection computed off-line. A power meter with multiple sensors could be configured to make simultaneous measurements and compute their ratio on the display of the meter.

Disadvantages here are that measurements are at a single frequency (not swept over frequency) and slow. This is what prompts the development of a more dedicated solution.



It is instructive to show how a simple reflectometer was made, as this is the basis of the scalar network analyzer. It was usual to use the incident signal to provide feedback for a level control on the source. The oscilloscope does not need much bandwidth but should have very high sensitivity. To calibrate the system a short (0dB return loss) would be used to provide a zero dB reference on the screen, this would be a wavy line which is traced with a marker pencil. The precision attenuator was set at 0dB for this line. By using the attenuator at 10 dB steps a family of wavy lines was marked on the screen. The attenuator was returned to 0dB and the device measured. Frequency was indicated by a "dip" in the trace caused by a wavemeter.

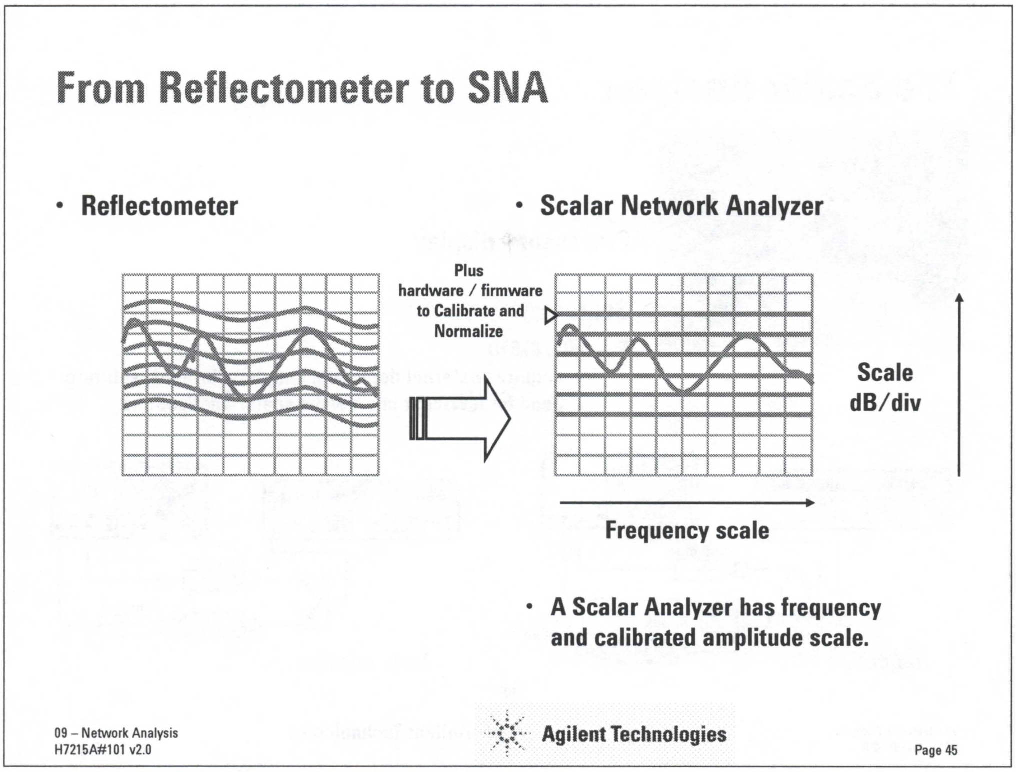
Reflectometer Display



09 – Network Analysis
H7215A#101 v2.0

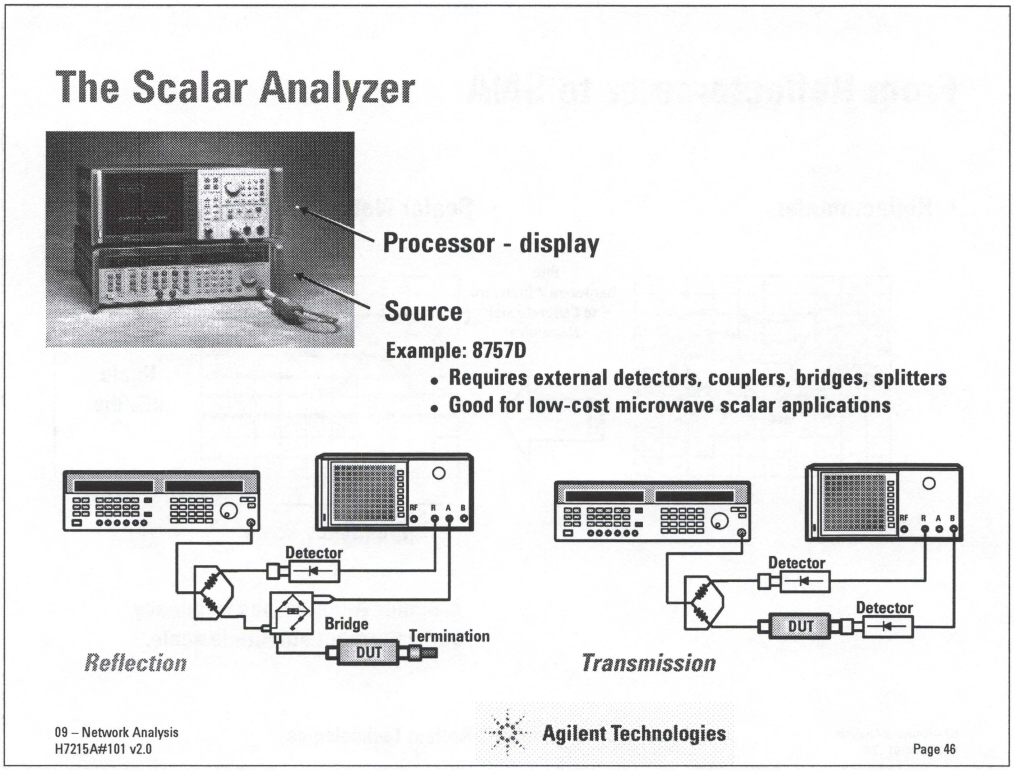
Agilent Technologies

Page 44



By adding features to the scope, such as trace normalization and a dB calibrated display, the reflection and transmission coefficient ratios may be computed and displayed directly.

This is essentially the evolution of the traditional scalar analyzer.

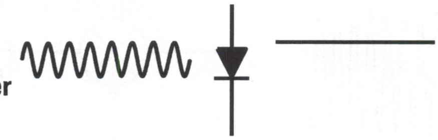


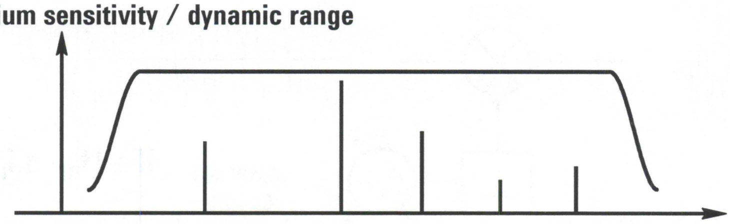
Here is a picture of a traditional scalar system consisting of a processor/display unit and a stand-alone source. This type of system requires external splitters, couplers, detectors, and bridges. Designed as a system, the scalar analyzer is maximized for speed and ease of use and incorporates many accuracy enhancement and convenience features. While not as common as they used to be, scalar systems such as this are good for low-cost microwave scalar applications.

The configuration shown for reflection provides the best measurement accuracy (assuming the termination is a high-quality Zo load), especially for low-loss, bidirectional devices (i.e., devices that have low loss in both the forward and reverse directions). Alternately, the transmission detector can be connected to port two of the DUT, allowing both reflection and transmission measurements with a single setup. The drawback to this approach is that the detector match (which is considerably worse than a good load) will cause mismatch errors during reflection measurements.

Broadband Diode Detection

- Easy to make broadband
- Inexpensive compared to tuned receiver
- Good for measuring frequency-translating devices
- Improve dynamic range by increasing power
- Medium sensitivity / dynamic range

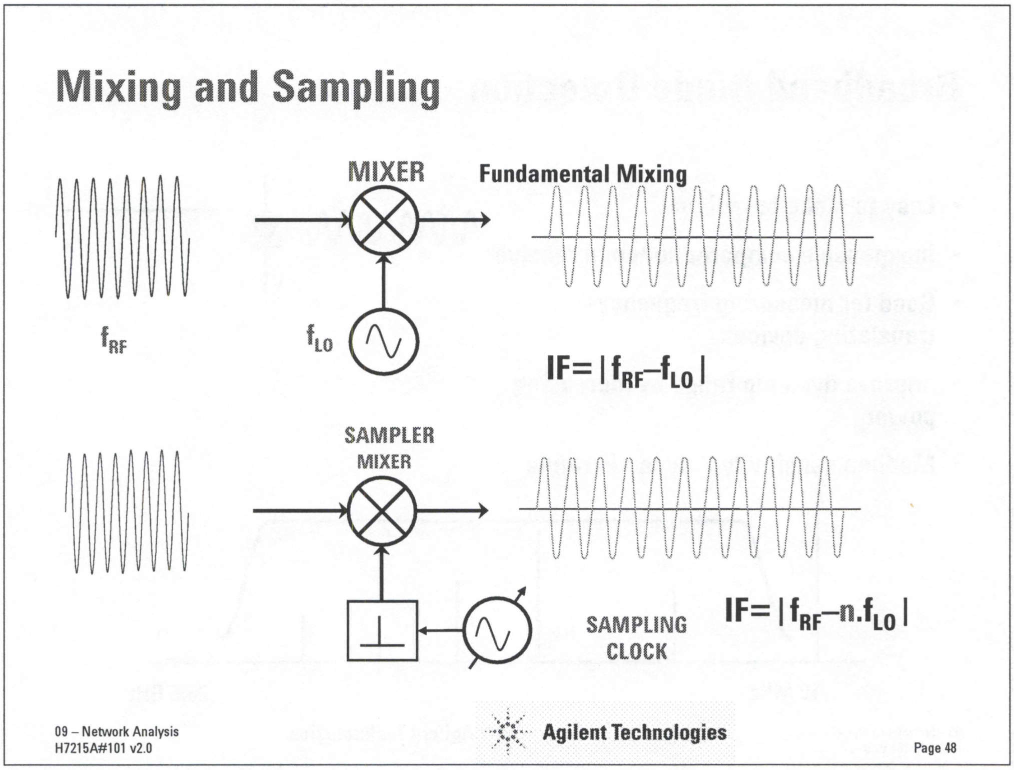




The main advantages of diode detectors are broadband frequency coverage (< 10 MHz on the low end to > 26.5 GHz at the high end) and they are inexpensive compared to a tuned receiver. Diode detectors provide medium sensitivity and dynamic range: they can measure signals to -60 dBm or so and have a dynamic range around 60 to 75 dB, depending on the detector type. Their broadband nature limits their sensitivity and makes them sensitive to source harmonics and other spurious signals. Dynamic range is improved in measurements by increasing input power.

AC detection eliminates the DC drift of the diode as an error source, resulting in more accurate measurements. This scheme also reduces noise and other unwanted signals. The major benefit of DC detection is that there is no modulation of the RF signal, which can have adverse effects on the measurement of some devices. Examples include amplifiers with AGC or large DC gain, and narrowband filters.

One application where broadband diode detectors are very useful is measuring frequency-translating devices, particularly those with internal LOs.

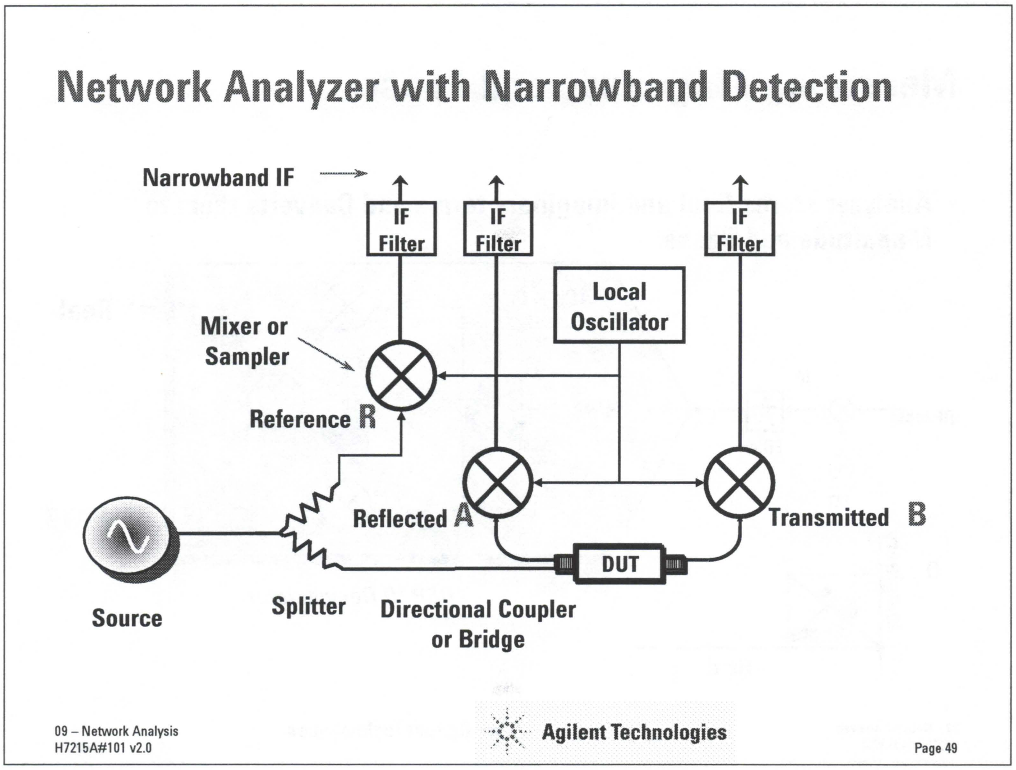


To solve some of the problems found in broadband detection, narrowband receiver techniques are used. By downconverting the wideband test signals to a lower narrowband frequency, signal processing techniques (filtering, amplification, magnitude and phase detection or A to D conversion) are more practical and economical. This technique is used in nearly all hetrodyne or narrow band receivers (including Network Analyzer Receivers).

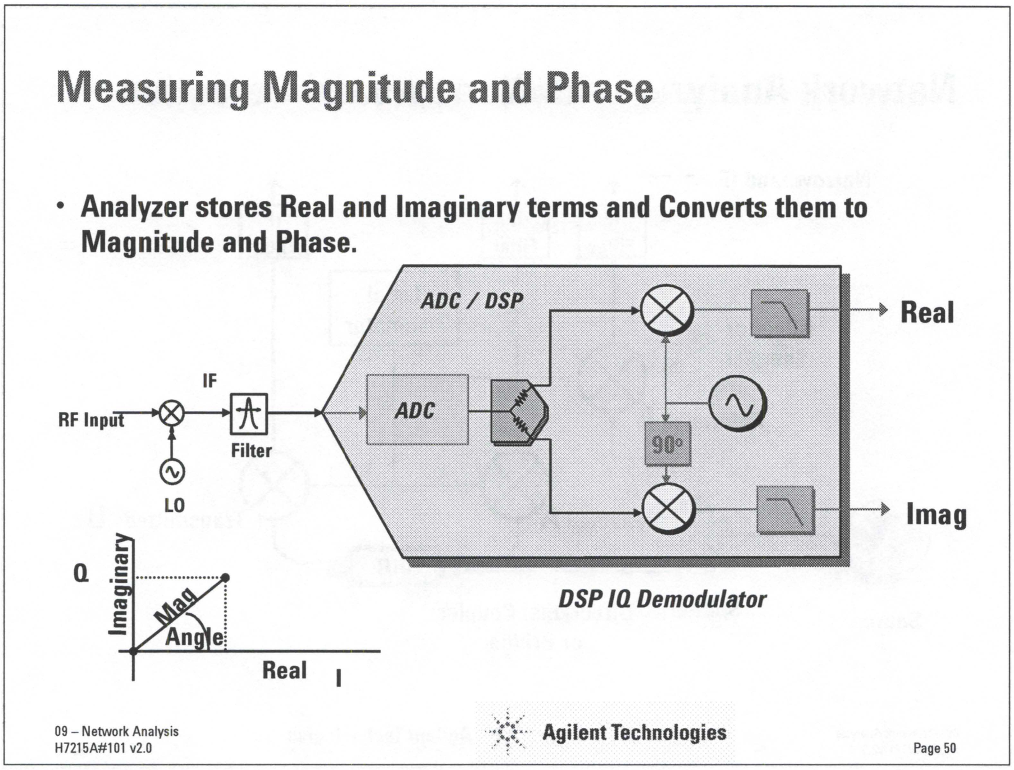
For microwave down conversion, sampling is a technique is often used. Very much like a mixer, the sampler is a diode-based device that is turned on and off, like a switch, by a sampling clock. During its "on" period, a sample of the RF waveform is presented to its output. If the sampling clock rate is chosen carefully, an undersampled, or aliased version of the RF signal is formed at the desired IF frequency. If the same LO is used to drive several mixers or samplers, the resulting IF signals will maintain the same magnitude and phase relationships as at RF.

Fundamental mixing is used in some receivers. This technique is expensive because the LO must have the same frequency range as the source but it is used for applications where high dynamic range is needed, or where the measurement down conversion is remote from the main analyzer, e.g. antenna testing.

There are lots of benefits to narrowband detection; IF filtering ensures only one signal measured at a time, filtering also reduces noise and increases measurement range and sensitivity, and more sophisticated processing allows the measurement of magnitude and phase.



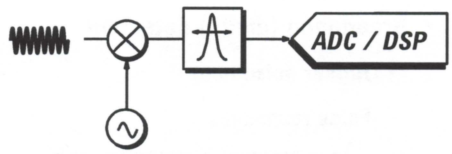
Here we see our reflectometer fitted with a true heterodyne receiver. One common LO or sampling clock ensures that the phase and magnitude of all channels are maintained. This is sometimes called a synchronous receiver. The IF filters pass only one mixing product and reduce system bandwidth and noise and prepare the signals to be further processed and finally detected.

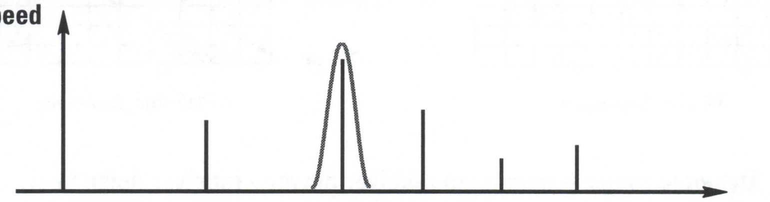


After the RF/microwave signal is downconverted and filtered, the analog signal is digitized by the Analog to Digital Converter (ADC). This digital information is passed to the Digital Signal Processor (DSP) which performs a demodulation of the complex signal into its Real and Imaginary parts. Separation of the complex signal is done with a In-phase/Quadrature-phase (IQ) demodulator. It doesn't matter that the IQ demodulator works on a digitized signal, the process of demodulation is the same whether an analog circuit or a digital signal processor is used.

Narrowband Detection - Tuned Receiver

- **Best sensitivity / dynamic range**
- **Provides harmonic / spurious signal rejection**
- **Improve dynamic range by increasing power, decreasing IF bandwidth, or averaging**
- **Trade off noise floor and measurement speed**





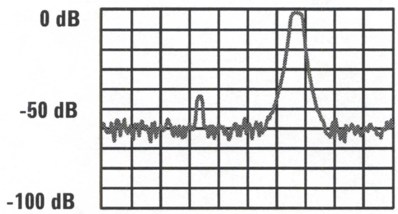
09 – Network Analysis
H7215A#101 v2.0
Agilent Technologies
Page 51

Tuned receivers provide the best sensitivity and dynamic range, and also provide harmonic and spurious-signal rejection. The narrow IF filter produces a considerably lower noise floor, resulting in a significant sensitivity improvement. For example, a microwave vector network analyzer (using a tuned receiver) might have a 3 kHz IF bandwidth, where a scalar analyzer's diode detector noise bandwidth might be 26.5 GHz. Measurement dynamic range is improved with tuned receivers by increasing input power, by decreasing IF bandwidth, or by averaging. The latter two techniques provide a trade off between noise floor and measurement speed. Averaging reduces the noise floor of the network analyzer (as opposed to just reducing the noise excursions as happens when averaging spectrum analyzer data) because we are averaging complex data. Without phase information, averaging does not improve analyzer sensitivity.

The same block diagram features that produce increased dynamic range also eliminate harmonic and spurious responses. As was mentioned earlier, the RF signal is downconverted and filtered before it is measured. The harmonics associated with the source are also downconverted, but they appear at frequencies outside the IF bandwidth and are therefore removed by filtering.

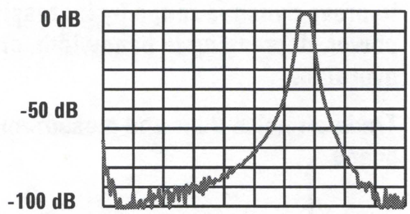
Comparison of Receiver Techniques

- **Broadband (diode) detection**
  - **Higher noise floor**
  - **False responses**



-60 dBm Sensitivity

- **Narrowband (tuned- receiver) detection**
  - **High dynamic range**
  - **Harmonic immunity**



< -100 dBm Sensitivity

- **Dynamic range = maximum receiver power - receiver noise floor.**

09 – Network Analysis
H7215A#101 v2.0

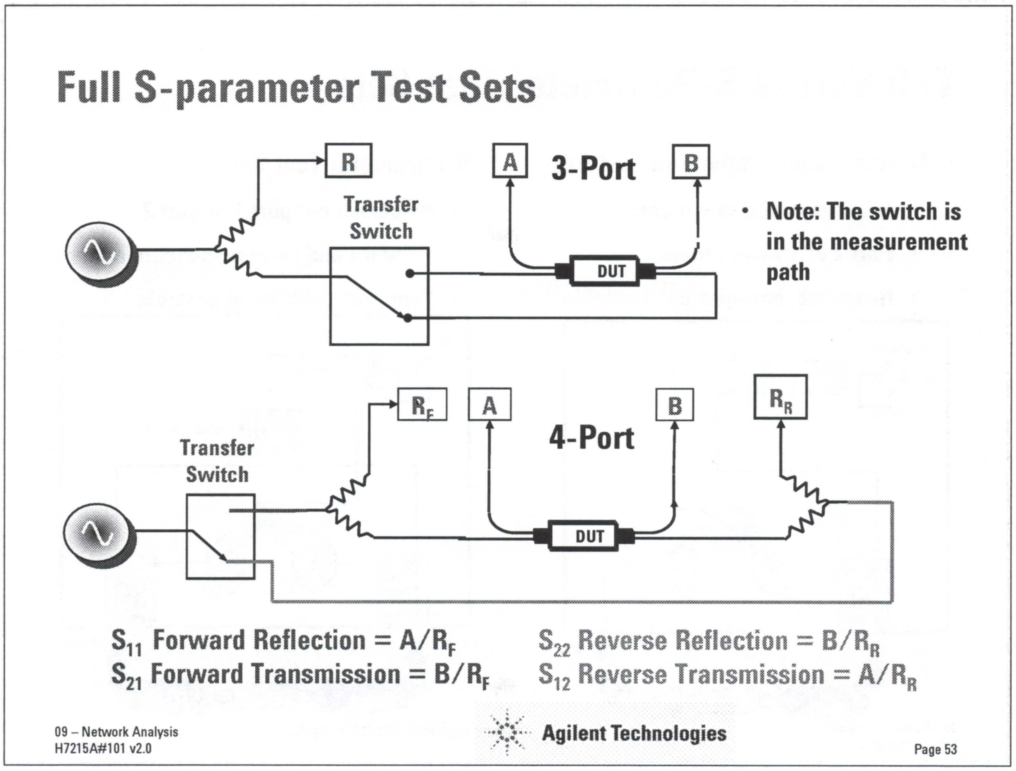
Agilent Technologies

Page 52

Dynamic range is generally defined as the maximum power the receiver can accurately measure minus the receiver noise floor. There are many applications requiring large dynamic range. One of the most common are filter applications. As you can see here, at least 80 dB dynamic range is needed to properly characterize the rejection characteristics of this filter. The plots show a typical narrowband filter measured on an 8757 scalar network analyzer and on the 8510 vector network analyzer. Notice that the filter exhibits 90 dB of rejection but the scalar analyzer is unable to measure it because of its higher noise floor.

In the case where the scalar network analyzer was used with broadband diode detection, a harmonic or subharmonic from the source created a "false" response. For example, at some point on a broadband sweep, the second harmonic of the source might fall within the passband of the filter. If this occurs, the detector will register a response, even though the stopband of the filter is severely attenuating the frequency of the fundamental. This response from the second harmonic would show on the display at the frequency of the fundamental. On the tuned receiver, a spurious response such as this would be filtered away and would not appear on the display.

Full S-parameter Test Sets



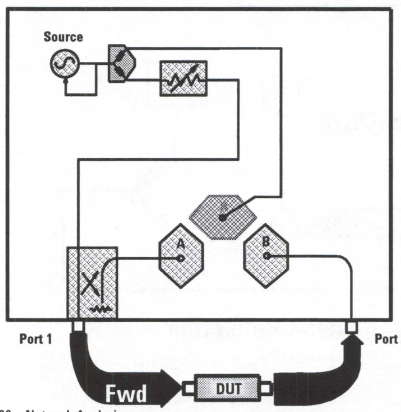
The reflectometer seen so far is often referred to as a transmission/ reflection test set. Notice that this basic T/R configuration will only measure the device in one direction. To obtain the reverse s-parameters, the device must be physically turned around. To improve upon this, the full s-parameter test set was developed. By adding a switch and some other components, the device can be reversed without removal from the circuit.

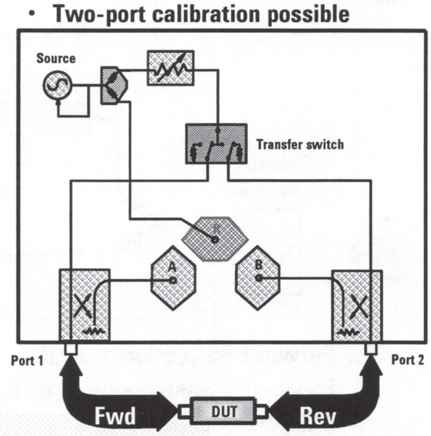
The upper illustration shows a low cost method of implementing a full s-parameter test set. Here, all that was required was the switch and an extra directional coupler. The disadvantage of this method is that the *switch is in the test path*. Any non-repeatability of the switch will become an error in the measurement.

The bottom illustration shows the addition of a switch, splitter, coupler and an extra measurement port. This is more expensive, but because the switch is located before the splitter, any instabilities due to the switch will be in both the reference and test paths and therefore ratioed out of the measurement. This also makes the test set symmetrical which has some other significant benefits.

T/R Versus S-Parameter Test Sets

- **Transmission/Reflection Test Set**
  - **RF always comes out port 1**
  - **Port 2 is always receiver**
  - **Response, one-port cal available**
- **S-Parameter Test Set**
  - **RF comes out port 1 or port 2**
  - **Forward and reverse measurements**
  - **Two-port calibration possible**





09 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 54

Here is a summary of the two basic types of test sets that are used with network analyzers. For transmission/reflection (T/R) test sets, the RF power always comes out of test port one and test port two is always connected to a receiver in the analyzer. To measure reverse transmission or output reflection of the DUT, we must disconnect it, turn it around, and re-connect it to the analyzer. T/R-based network analyzers offer only response and one-port calibrations, so measurement accuracy is not as good as that which can be achieved with S-parameter test sets. However, T/R-based analyzers are more economical.

S-parameter test sets allow both forward and reverse measurements on the DUT, which are needed to characterize all four S-parameters. RF power can come out of either test port one or two, and either test port can be connected to a receiver. S-parameter test sets also allow full two-port (12-term) error correction, which is the most accurate form available. S-parameter network analyzers provide more performance than T/R-based analyzers, but cost more due to extra RF components in the test set.

RF & Microwave Measurement Fundamentals

10 — Network Analysis

Agilent Technologies

General Measurement Procedure

1. Warm-up the Analyzer
2. Clean, inspect and gauge all connectors and cables
3. Connect cables and adapters to the analyzer
4. PRESET the analyzer
5. Setup stimulus frequencies and power
6. Connect the DUT to verify setup and operation
7. Select S-parameter(s) to be measured
8. Select the display Format
9. Scale the display
10. Remove DUT
11. Calibrate
12. Verify calibration
13. Re-connect DUT
14. Save the instrument state

10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 2

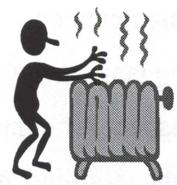
This procedure is a typical procedure for making any measurement with a network analyzer. Considering each of these steps carefully will help to insure accurate results.

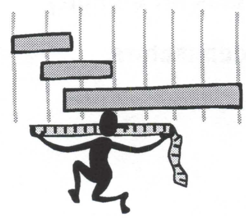
Unlike most instruments, the network analyzer arrives unconfigured and uncalibrated for measurements. Because of the high accuracy required for this work, the analyzer must be set up and calibrated by the user for each unique device under test.

Let's discuss the major steps of the process.

General Measurement Procedure

- **PREPARE**
  - **Warm-up the Analyzer**
  - **Check connectors and cables**
  - **Connect cables and adapters to the analyzer**
  - **Connect the DUT to verify setup and operation**
- **CALIBRATE**
  - **Remove DUT**
  - **Calibrate**
  - **Verify calibration**
  - **Save the instrument state**
- **MEASURE**
  - **Re-connect DUT**
  - **Measure**





10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 3

This procedure is a typical procedure for making any measurement with a network analyzer. Considering each of these steps carefully will help to insure accurate results.

Unlike most instruments, the network analyzer arrives unconfigured and uncalibrated for measurements. Because of the high accuracy required for this work, the analyzer must be set up and calibrated by the user for each unique device under test.

Let's discuss the major steps of the process.

Clean, Inspect and Gauge All Cables & Connectors

- **Use precision connectors wherever possible**
- **Clean using isopropyl alcohol & lint-free swabs/patches**
- **Visually inspect for damage**
- **Use only phase-stable, high quality cables**
- **Test cables for stability**
- **Gauge connectors**

10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 4

After the analyzer and the associated hardware have attained stabilized at their operating temperature, the hardware can be inspected and prepared for use. In spite of the capability of the analyzer to compensate for hardware imperfections, this capacity can only be stretched so far and will not compensate for poor and unstable adapters and cables. Therefore, the highest quality connectors, adapters and cables must be used. Only then can we compensate for the more subtle and repeatable imperfections.

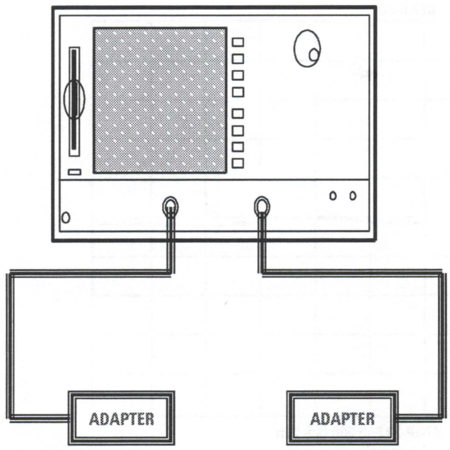
All connectors must be clean. Modern analyzers can "measure" a single fingerprint. Allowing dirt to remain will not only make measurements inaccurate and unrepeatable but will ruin the expensive cables and calibration kits.

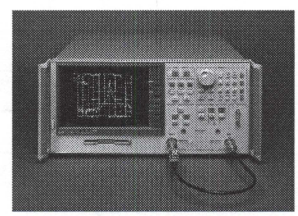
Careful visual inspection of all connectors before each use is probably the most important step to avoid measurement problems and propagating damage from connector to connector.

Cables are the greatest source of instability in the system. Phase shifting due to cable flexure is very hard to control without using cables specifically designed for network analyzer use. Once connected, flex the cables and look carefully at the measurements for signs of cable and other mechanical instabilities.

On a regular basis, all system connectors should be checked for mechanical tolerance by using a set of specialized gauges and micrometers.

Connect Cables and Adapters





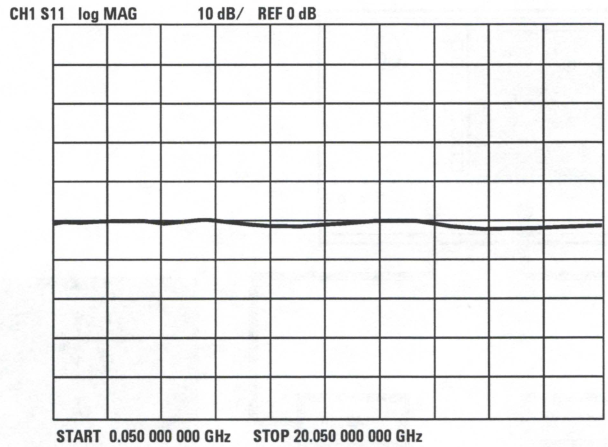
10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 5

After the inspection of the hardware is complete, connect all of the cables and adapters required to link the analyzer to the device under test (DUT). These connections should be solid mechanical connections tightened with a torque wrench. There are specific torque wrenches and torques recommended for each family of connector. These connections must not be broken or even loosened again until after the measurements are complete.

PRESET the Analyzer



The network analyzer is a fairly complicated instrument with hundreds of menu choices and parameters that can be set by the operator. To verify that all of these are in a desirable state would take a long time. The best way to be sure that the analyzer is in a benign state is to press the preset button and all functions will be set to a simple default state. This state is usually a measurement of S11 forward reflection, a scale of 10 dB per division, reference line located at the center of the grid and set to a value of 0 dB. The frequency span at preset is the full range of the analyzer, and functions like error correction, and other offsets are off.

Setup Stimulus

- **Adjust start/stop, center/span frequencies**
- **Select number of points**
- **Set source power**
  - **High for best dynamic range**
  - **Specific for active devices**

10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 7

The stimulus parameters should be setup early in any network analyzer measurement. S parameter measurements and their associated error corrections have magnitudes and phases which are unique to each frequency and power level.

Start and stop or center and span frequencies may be set using the hard keys found in the stimulus block of the front panel. Choose the frequency span based upon the operating parameters of the DUT and the required measurement results. When the DUT is not well understood, connect the DUT and set the frequency span based upon observations made on the CRT.

Other parameters like power level and number of frequency points can be set under the stimulus menu. The default number of frequencies is 201 points evenly distributed between the start and stop frequencies. The number of points may be adjusted between 3 and 1601. It is also possible to create a table of specific frequencies for test. Any calibrations that are made are unique to the stimulus frequency set. If the frequencies are changed after calibration, the error corrections will be turned off!

The power of the internal synthesized source can also be adjusted. For devices that are not sensitive to RF power levels, keep the power level high (0 dBm and above) for maximum measurement dynamic range. For devices like amplifiers and transistors, be careful to choose the proper power level for test and to not overdrive the component. In situations like this, it is recommended to verify the power at the DUT with a power meter. There is also a power meter calibration feature to insure a specific power level at the DUT.

Connect the DUT to verify setup and operation

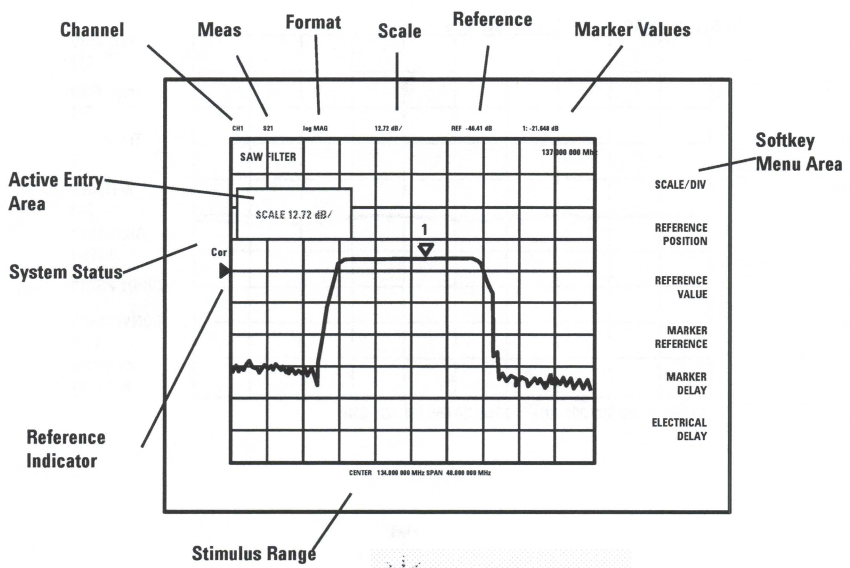
- **Confirm choice of adapters**
- **Check for mechanical stability**
- **Check DUT for proper operation**
- **Saves unnecessary calibrations**

A step that may seem unnecessary is to connect the DUT into its place in the measurement system. The network analyzer has not yet been calibrated and will not yield high accuracy results, but developing the habit of performing this check will save time and money in the long run.

Making this connection will confirm that all of the cables and adapters will accommodate the DUT. It is not unusual to discover that the cables are not long enough, the DUT cannot be supported, or the connectors are of the wrong type or sex. If we were to go ahead and calibrate the system with these problems in place, calibration would need to be repeated after the problems were corrected wasting time and causing unnecessary wear on the calibration kit. Each hardware setup and frequency set must be calibrated separately.

Once the DUT is in place, check it for proper operation. There is no need to proceed with a calibration for a DUT that is mistuned of malfunctioning.

Interpreting the Display



10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 9

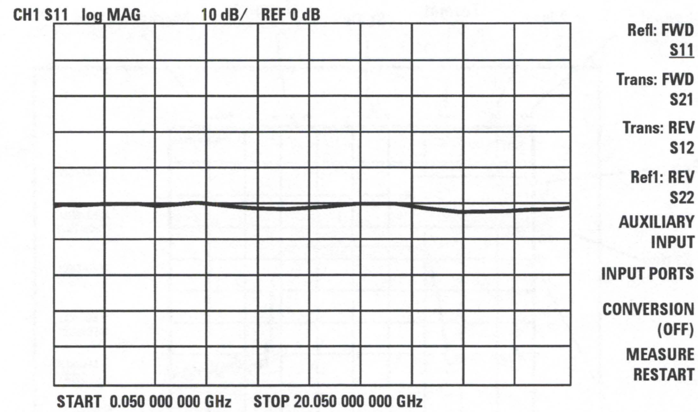
Operation of the DUT may be confirmed by performing the next few steps in the general measurement procedure and interpreting the results on the display.

Looking along the top of the measurement grid from left to right, the s-parameter selected may be seen followed by its display format (log mag, swr, polar, etc.). Next the scale per division and reference line information are noted. The reference line is an additional horizontal graticule that can be positioned anywhere on the grid and assigned any value in the current format as a convenient reference point. If a marker is turned on, its position in both axis are shown to the upper right of the grid.

Below the grid are the stimulus parameters of start/stop or center/span. These will be in Hertz in the frequency domain and seconds in the time domain.

The active entry area notes the value of the analyzer setting being adjusted. To the left of the grid is a system status area where notes the operator is apprized of the state of error correction, averaging, smoothing, certain offsets, and error conditions.

Select S-parameter to be measured



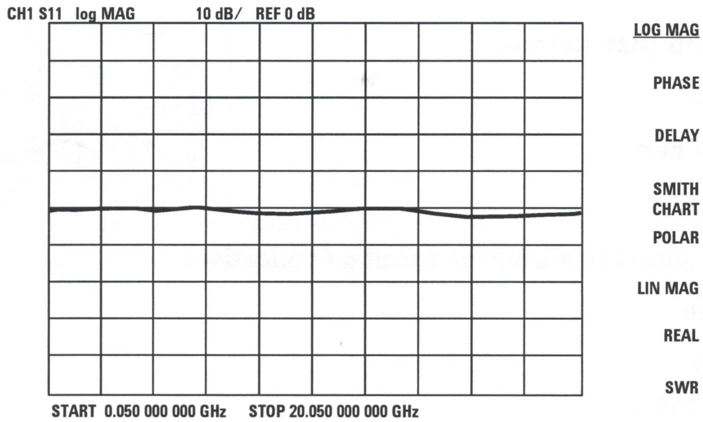
10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 10

The s-parameter to be measured can be selected from the Measurement menu.

Select the Display FORMAT



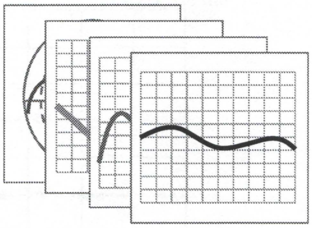
10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 11

Use the Format key to select the display format.

Common Display Formats for S11 (& S22)

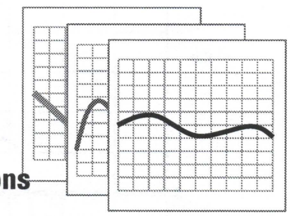


- **Log MAG (Return Loss)**
- **SWR**
- **Smith Chart**
- **Polar**
- **Other Formats Available for Specific Applications**
  - **Phase**
  - **Delay**
  - **Lin MAG**
  - **Real**
  - **Imaginary**

10 – Network Analysis
H7215A#101 v2.0
Agilent Technologies
Page 12

There are many display formats available to the user when measuring S_{11} and S_{22}, but generally only a few are widely used. They are the Log MAG, SWR, Smith chart and POLAR formats. The other format options can be used but generally have very specific measurement applications. For example, the Linear MAG format is a unit-less display of the reflection coefficient magnitude and is useful for conversion parameters and time domain data.

Common Display Formats for S21 (& S12)



- Log MAG (Insertion Loss)
- Phase
- Delay
- Other Formats Available for Specific Applications
  - Smith Chart
  - Polar
  - Lin MAG
  - SWR (not valid for S21, S12)
  - Real
  - Imaginary

10 – Network Analysis
H7215A#101 v2.0

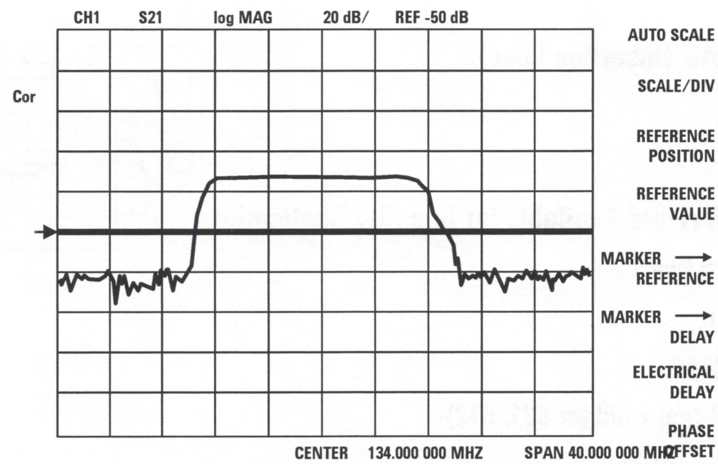
Agilent Technologies

Page 13

There are many display formats available to the user when measuring the transmission response (S_{21} and S_{12}) but generally only a few are widely used. They are the Log MAG, PHASE (with electrical delay) and DELAY formats. The other format options can be used but generally have very specific measurement applications. For example, the Linear MAG format is a unit-less display of the transmission coefficient magnitude and is useful when measuring the time domain response of a DUT. Note, the SWR format has no valid use when measuring the transmission through the DUT.

We will discuss the more advanced formats and their use later in the course.

Scale the Display



10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

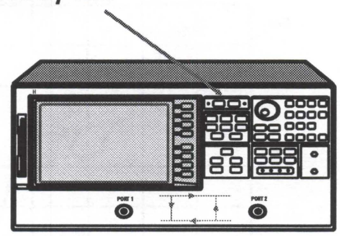
Page 14

The Scale menu allows the operator to scale the grid in its current format and adjust the position and value of the reference line. The autoscale function will scan the measurement trace array and select a scale per division and a reference line value that will fit the trace to the grid. The Scale key allows the operator continuous adjustment of the scale per division. The scale of a network analyzer is not confined to specific scales like many other instruments. Likewise, the reference position and value may be continuously adjusted anywhere on the grid and to a wide range of values.

The scale menu also allows some offsets to be added to the measurement values.

Measurement Channels

- **Channels 1 and 2 can be set-up independently for:**
  - **S-parameter**
  - **Display format**
  - **Scale and Reference level**
  - **Domains**
- **And by uncoupling the channels:**
  - **Stimulus frequency and power parameters**
  - **Calibration**



10 – Network Analysis
H7215A#101 v2.0

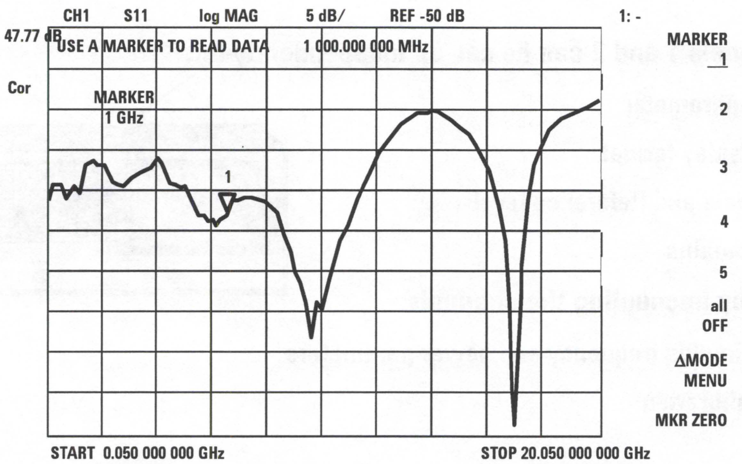
Agilent Technologies

Page 15

For greater flexibility, these analyzers have 2 measurement "channels". A channel in this case is not a physical port or connection point but a measurement setup. Each of channels 1 and 2 may be setup independently for s-parameter, display format, scale and reference level. This allows the operator to quickly switch between the two setups such as reflection and transmission measurement on the same DUT. These two channels can be viewed separately or together on the display.

In the default instrument configuration, the channels share the stimulus parameters and error correction. By uncoupling the channels under the stimulus menu, the two channels may also have independent frequency sets, stimulus power and error correction. This allows the DUT to be evaluated under different conditions and spans like a wide span and a narrow span viewed on the same display.

Press MKR, Try the Knob



10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 16

Now that the display is configured in some useful way, measurements can be easily made using markers. Simply press the Marker button to turn on a marker. The position and value read by the marker is shown in the upper right-hand corner of the grid. The location of the marker can be changed by turning the knob or typing a frequency directly on the keypad.

General Measurement Procedure -Remember These?

1. Warm-up the Analyzer
2. Clean, inspect and gauge all connectors and cables
3. Connect cables and adapters to the analyzer
4. PRESET the analyzer
5. Setup stimulus frequencies and power
6. Connect the DUT to verify setup and operation
7. Select S-parameter(s) to be measured
8. Select the display Format
9. Scale the display
10. Remove DUT
11. Calibrate
12. Verify calibration
13. Re-connect DUT
14. Save the instrument state

10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 17

This procedure is a typical procedure for making any measurement with a network analyzer. Considering each of these steps carefully will help to insure accurate results.

Unlike most instruments, the network analyzer arrives unconfigured and uncalibrated for measurements. Because of the high accuracy required for this work, the analyzer must be set up and calibrated by the user for each unique device under test.

Let's discuss the major steps of the process.

Lab

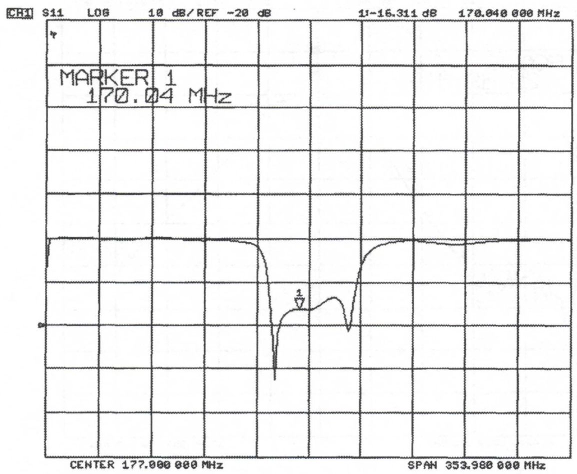
- Follow the General Measurement Procedure but omit calibration.
- Mission:
  - Set up the network analyzer to measure the DUT
    - Measure S11 in dB (return loss)
    - Measure S21 in dB (insertion loss)
    - Display both measurements simultaneously
- Need to Know:
  - Device Type (Filter, amplifier, etc.
  - Start/Stop or Center/Span Frequencies
  - Stimulus Power Limitations

10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 18

Lab Results

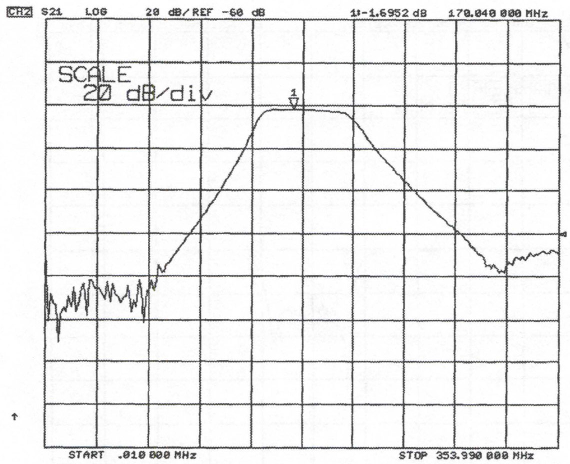


10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 19

LAB Results

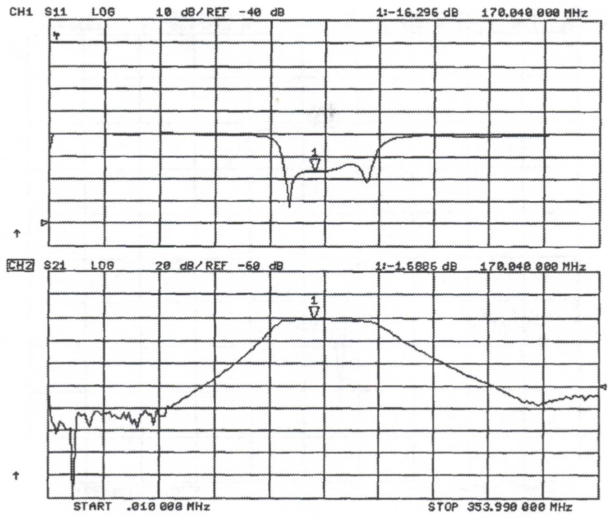


10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 20

Lab Results

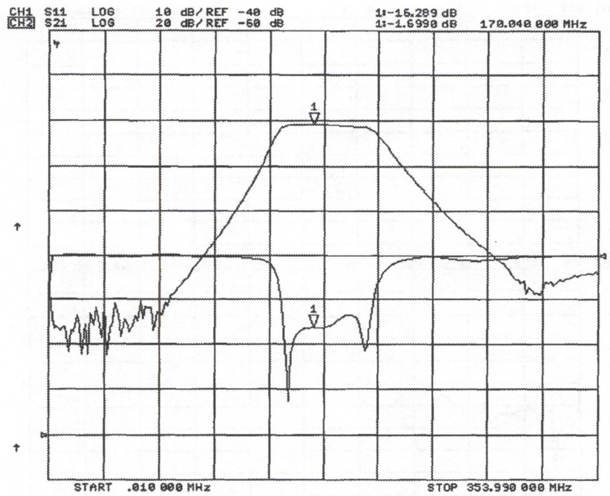


10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 21

Lab Results



10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 22

Sources of Error

10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 23

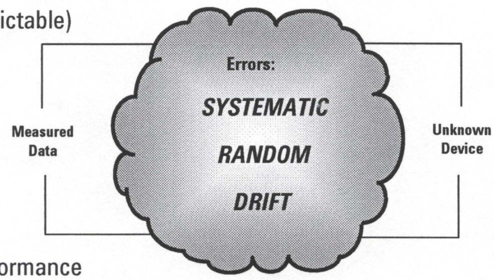
Measurement Error Modeling

Systematic errors

- due to **imperfections** in the analyzer and test setup
- are assumed to be **time invariant** (predictable)
- <u>can</u> be characterized (during calibration process) and **mathematically removed** during measurements

Random errors

- **vary** with time in random fashion (unpredictable)
- **<u>cannot be removed</u>** by calibration
- main contributors:
  - **instrument noise** (source phase noise, IF noise floor, etc.)
  - **switch** repeatability
  - **connector** repeatability



Drift errors

- are due to instrument or test-system performance changing ***after*** a calibration has been done
- are primarily caused by **temperature variation**
- **<u>can be removed</u>** by further calibration(s)

10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

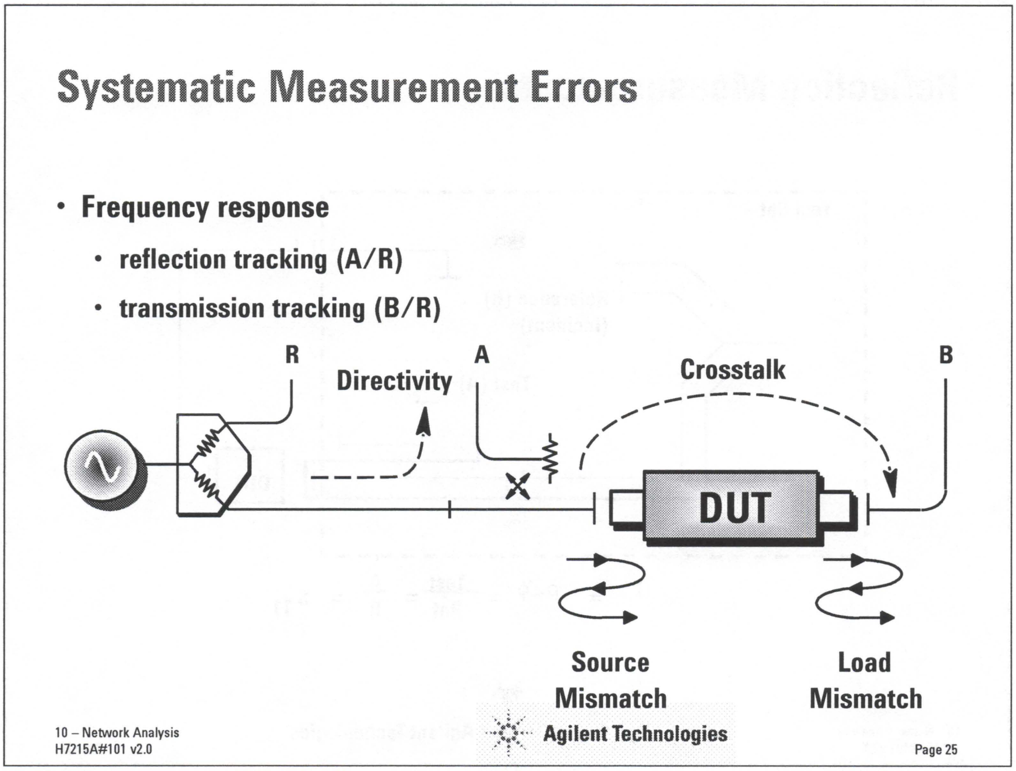
Page 24

Let us look at the three basic sources of measurement error: systematic, random and drift.

Systematic errors are due to imperfections in the analyzer and test setup. They are repeatable (therefore predictable), and assumed to be time invariant. Systematic errors are characterized during the calibration process and mathematically removed during measurements.

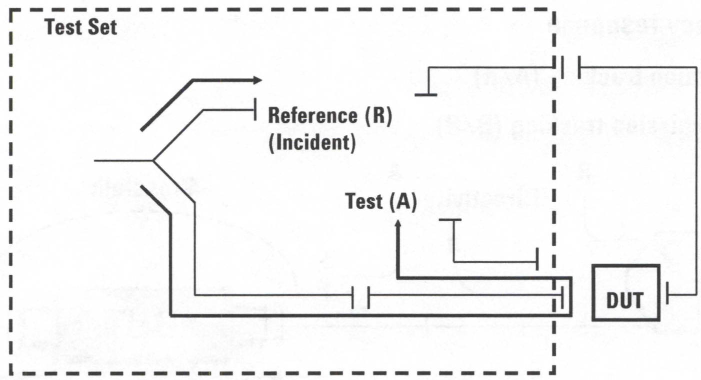
Random errors are unpredictable since they vary with time in a random fashion. Therefore, they cannot be removed by calibration. The main contributors to random error are instrument noise (source phase noise, sampler noise floor) and mechanical errors from cable flexure and non-repeatable connections.

Drift errors are due to the instrument or test-system performance changing after a calibration has been done. Drift is primarily caused by temperature variation and it can be removed by further calibration(s). The timeframe over which a calibration remains accurate is dependent on the rate of drift that the test system undergoes in the user's test environment. Providing a stable ambient temperature usually goes a long way towards minimizing drift.



Shown here are the major systematic errors associated with network measurements. The errors relating to signal leakage are directivity and crosstalk. Errors related to signal reflections are source and load match. The final class of errors are related to the frequency response of all of the hardware involved. All system components will have a finite frequency response. If the reference path and measurement path had identical frequency responses, these effects would ratio out in the final result. Unfortunately the paths are not identical and do not "track" each other perfectly. Therefore the differences in frequency response between the paths is often called reflection and transmission tracking errors.

Reflection Measurements



$$\Gamma = \rho\angle\phi = \frac{\text{Test}}{\text{Ref}} = \frac{A}{R} = S_{11}$$

10 – Network Analysis
H7215A#101 v2.0

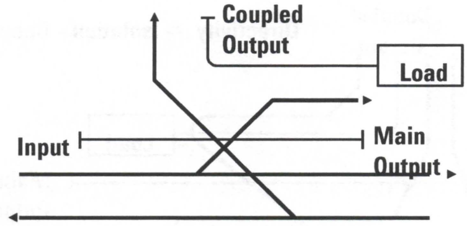
Agilent Technologies

Page 26

For measurement of the input reflection coefficient of a RF/microwave device, we are interested in determining the ratio of the signal reflected from the device to the incident signal (represented by the signal in the reference channel).

Directivity

- **Ideal coupler**
  - **Perfect separation of forward and reverse traveling waves**

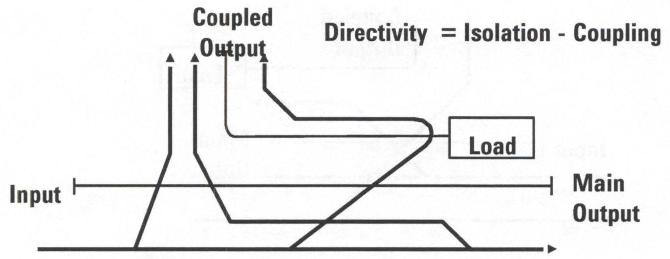


10 – Network Analysis
H7215A#101 v2.0
Agilent Technologies
Page 27

Normally, a device that can separate the reverse from the forward traveling waves (a coupler in this example), is used to detect the signal reflected from the device. Ideally, the coupler will completely separate the incident and reflected signals, and only the reflected signal appears at the coupled output.

Directivity (Continued)

- **Real coupler**
  - **Imperfect separation due to leakages and spurious reflections (directivity)**



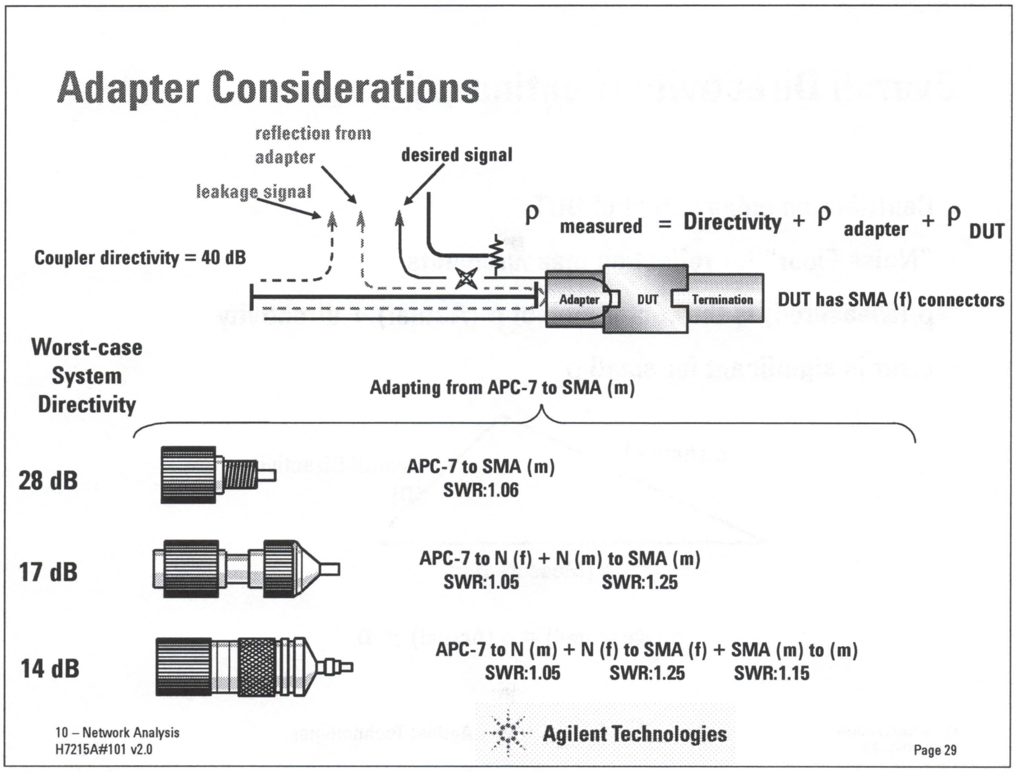
10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 28

However, a real coupler is not perfect. A small amount of the incident signal will appear at the coupled output due to leakage as well as reflection from the termination in the coupled arm. Also, reflections from the coupler output connector will appear at the coupled output, adding uncertainty to the signal reflected from the device.

The figure of merit for how well a coupler separates forward and reverse waves is directivity. Directivity is defined as the ratio of power coupled into the coupled arm when the coupler is in the forward direction to the power available in the coupled arm when the coupler is in the reverse direction, using a Zo termination.



Directivity of the network analyzer system includes not only stray reflections from inside the directional coupler but reflections from anything other than the DUT. Any cables or adapters that are placed between the DUT and the coupled output of the directional coupler will cause reflections that degrade the system directivity. Therefore, this additional hardware must be eliminated whenever possible and when required, must be of the highest quality.

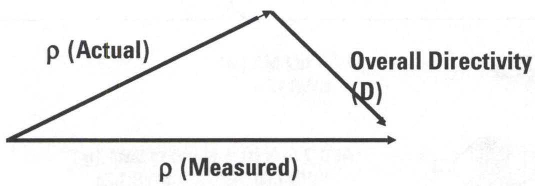
The slide shows the effect of adding even high-quality adapters with low SWR's (reflection) has a large effect on the total system directivity. The adapter causes an error signal which can add or subtract with the desired signal from the DUT.

If the adapter has a SWR of say 1.5 (the less-expensive variety), the effective directivity of the coupler drops to around 14 dB worst case, even if the coupler itself had infinite directivity! In other words, with a perfect Zo load on the output of the adapter; the reflected signal appearing at the coupled port would only be 14 dB less than the reflection from a short or open circuit. Stacking adapters compounds the problem, as is illustrated above. Consequently, it is very important to use quality adapters (or preferably, no adapters at all) in your measurement system, so system directivity is not excessively degraded. While error-correction can mitigate the effect of adapters on the test port, our system is more susceptible to drift with degraded raw (uncorrected) directivity.

Overall Directivity (Continued)

- **Contribution independent of DUT**
- **"Noise Floor" for reflection measurements**

  ρ (Measured) is the vector sum of ρ (Actual) + directivity
- **Error is significant for small ρ**



ρ (Measured) = ρ (Actual) ± D

10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 30

The uncertainty contributed by directivity is independent of the test device. This error signal is constantly at the output of the coupler and therefore represents the minimum reflection signal that can be measured. It can be thought of as the "noise floor" for reflection limiting the sensitivity of the measurement.

Like the signal reflected from the test device, the directivity signal is a vector quantity and the level of its contribution will depend on the phase relationship of the two signals.

Its most significant impact is in the measurement of devices with good match (small rho) and can obscure the true measurement completely.

Frequency Response (Tracking)

- **Vector sum of test setup variation in frequency response**
  - **Signal separation device**
  - **Cables/adapters**
  - **Test/reference channel variation**
- **Contribution dependent on device**

$$\text{Tracking Error} = \text{Tracking} \cdot \rho$$

10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 31

The frequency response of the measurement system will add to overall uncertainty. The system frequency response is the vector sum of all test setup variations in magnitude and phase frequency response, including signal separation device, test cables, adapter, and variations in frequency response between the reference and test channels. The error due to frequency response is dependent on the input reflection coefficient of the test device.

Source Match

- **Impedance mismatch at test device looking back into source**
- **Adapter/cable mismatches and losses**

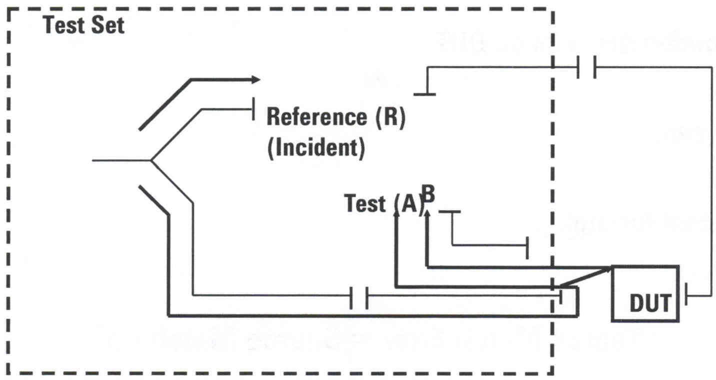
10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 32

A second major systematic error in reflection measurements is source match. Source match is defined as the vector sum of signals appearing at the system test input due to the impedance mismatch at the test device looking back into the source port, as well as adapter and cable mismatches and losses.

Source Match (Continued)



A $= \rho$
B $=$ Source Match $\cdot \rho^2$

10 – Network Analysis
H7215A#101 v2.0
Agilent Technologies
Page 33

As can be seen in the example, the desired signal (A) is not the only signal present at the system test input. Some of the signal reflected from the test device will reflect from the system port and re-reflect from the test device to produce a residual signal at the system test input (B).

Source Match (Continued)

- **Contribution depends on DUT**
- **Vector sum**
- **Significant for high ρ**

$$\textbf{Source Match Error} = \textbf{Source Match} \cdot \rho^2$$

10 – Network Analysis
H7215A#101 v2.0

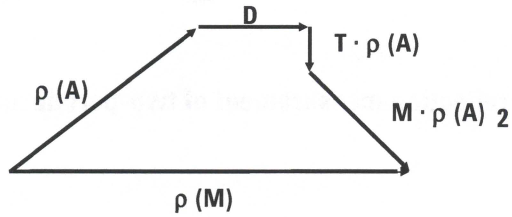
Agilent Technologies

Page 34

The uncertainty contributed by source match is dependent on the characteristics of the test device and adds to the measurement in a vector fashion. The level of uncertainty contributed is most significant for devices with a high input reflection coefficient.

Reflection Measurement Error Analysis

- **Vector sum**



- **Worst case**

$$\rho(M) = \rho(A) \pm D \pm T \cdot \rho(A) \pm M \cdot \rho(A)^2$$

10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 35

In the final result, the uncertainty contributed by these primary sources of error in a reflection measurement will be the vector sum of the individual errors. There are numerous methods for determining what the overall uncertainty in a measurement are. Worst case addition will give the highest confidence level, as well as the most pessimistic projection. This analysis assumes a 1-port device or a device with high transmission isolation between the input and output ports.

Load Match

- **Imperfect termination of DUT output**
- **Factor in reflection measurement of two-port devices**

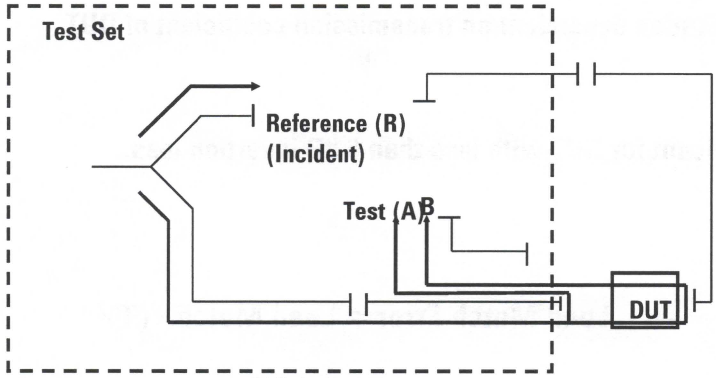
10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 36

One other potential source of error in a reflection measurement exists when measuring the input reflection coefficient of a two-port device. It results from the imperfect termination of the device output port by the test system.

Load Match (Continued)



$A = \rho$

$B = \text{Load Match} \bullet (\text{tau})^2$

10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 37

As can be seen in this example, some of the energy transmitted through the test device will reflect from the return port and add energy to the signal detected at the system test input. The transmission coefficient (tau) is encountered twice by the error signal.

Load Match (Continued)

- **Contribution dependent on transmission coefficient of DUT**
- **Significant for DUT with less than 6 dB insertion loss.**

$$\textbf{Load Match Error} = \textbf{Load Match} \cdot (T)^2$$

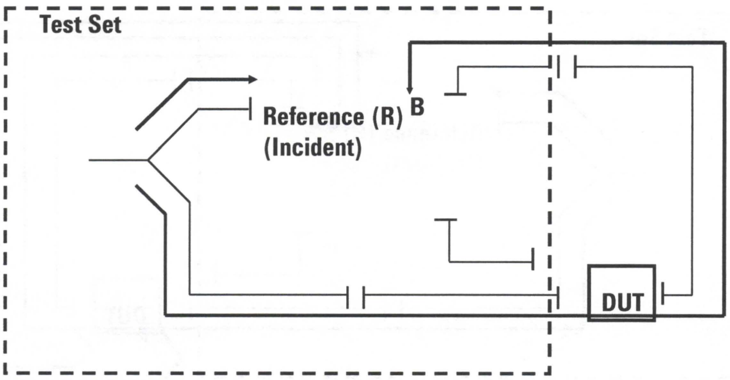
10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 38

This source of error is also commonly known as the "two-port effect." The contribution to error of this effect is very dependent on the transmission coefficient of the test device and is important to consider when measuring a bilateral device. Its impact is usually ignored when the test device insertion loss is greater than 6 dB that will provide greater than 12 dB isolation between ports. This error will cause large uncertainties when making a reflection measurement of a transmission line such as a cable, coupler or microstrip device.

Transmission Measurements



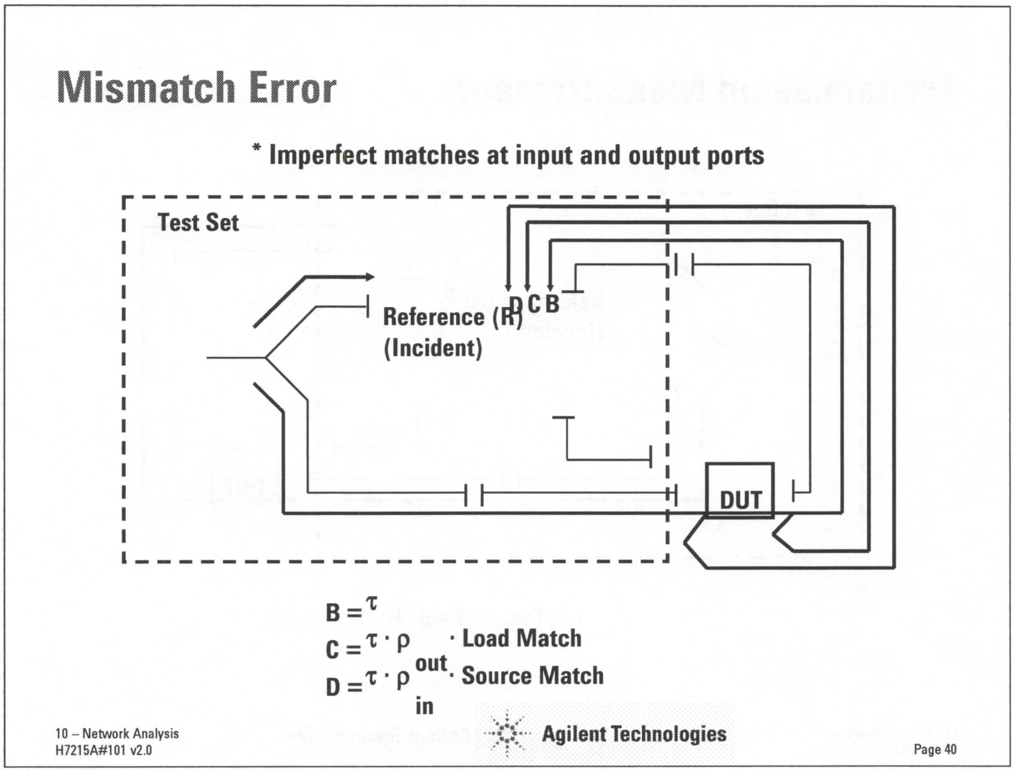
τ = Test/Ref =B/R= S_{21}

10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 39

For transmission measurements, the ratio of interest is that of the energy transmitted through the device to the incident energy. The systematic errors encountered in a transmission measurement are similar to those present in a reflection measurement.



Mismatch errors also affect transmission measurements. As the transmitted signal encounters mismatches between the network analyzer ports and the DUT, these reflections eventually show up at the test port as a vector sum with the desired transmitted signal. The slide shows the simple case of reflections caused by source and load match but there can be many permutations of these reflections.

Mismatch Error (Continued)

- Contribution dependent on DUT input/output ρ
- Significant for DUT with low t and high input or output ρ

$$\text{Load match error} = \tau \cdot \rho_{out} \cdot \text{Load Match}$$

$$\text{Source match error} = \tau \cdot \rho_{in} \cdot \text{Source Match}$$

10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 41

The contribution is heavily dependent on the test device input and output reflection coefficients. Poorly matched devices will suffer more from mismatch errors.

If the DUT has a low transmission coefficient (low loss), the reflections are permitted to propagate back and fourth through the device un-attenuated. The measurement of many common devices such as filters and microstrip devices will be complicated due to their reflective input and output ports and their low loss characteristics.

Isolation (Crosstalk)

- **Leakage between test and reference channels**
- **Includes RF and IF signal path**

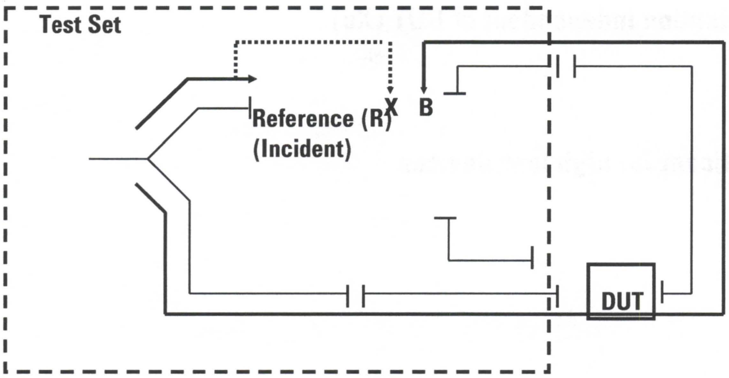
10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 42

Leakage of energy between the system test and reference channels contributes to error in a transmission measurement much like directivity does to a reflection measurement. This is energy that was not transmitted through the DUT but found an alternate path to the test channel. These paths may be found inside the test set, outside the test set between cables and in fixturing or even in the IF processing sections of the network analyzer. Energy from any of these will limit the measurement range of the analyzer for transmission measurements.

Isolation (Crosstalk) (Continued)



B = τ
X = Crosstalk

10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 43

Crosstalk is shown here as a leakage into the the test channel

Isolation (Crosstalk) (Continued)

- **Contribution independent of DUT (Xc)**
- **Significant for high loss devices**

10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 44

Like directivity, isolation error is independent of the characteristics of the devices under test. This error is significant for devices with high insertion loss (>60 dB) such as the rejection band of a filter.

Frequency Response (Tracking)

- **Vector sum of test setup variation in frequency response**
  - **Cables/adapters**
  - **Test/reference channel variation**

- **Contribution dependent on device**

Transmission tracking error = Tracking $\cdot$ τ

10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

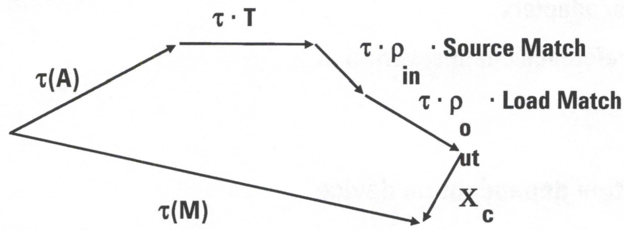
Page 45

The final error to be discussed is frequency response. All hardware components; cables, couplers, splitters, etc. have a frequency response. If the frequency response of the reference and test path were the same, these errors would normalize or "track" out in the ratioing process. However, these two paths are different so we must account for the differences in frequency response between the paths. This is why this error is referred to as a "tracking error".

The DUT is in the test path for a transmission measurement so therefore the tracking error is dependent upon the transmission coefficient of the DUT.

Transmission Measurement Error Analysis

- Vector sum



- Worst case

$$\tau(M) = \tau(A) \pm \tau \cdot T \pm \tau \cdot \rho_{in} \text{ Source Match} \pm \tau \cdot \rho_{out} \cdot \text{Load Match}$$

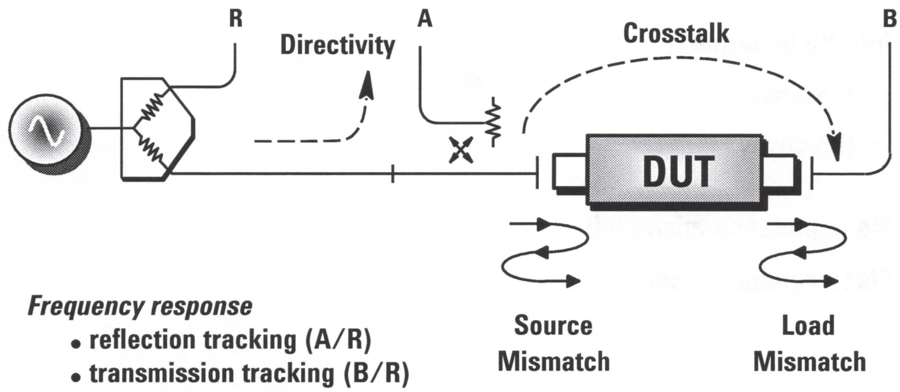
10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 46

In the final analysis, the total uncertainty is the vector sum of all individual error sources. The magnitude and phase of each of these errors will vary at each frequency and so will the vector sum. Therefore all of these errors must be managed at each stimulus frequency. When sweeping many frequencies over a wide bandwidth, it is highly likely that the worst case sum of these errors will be encountered.

Systematic Measurement Errors



Frequency response
- **reflection tracking (A/R)**
- **transmission tracking (B/R)**

Six forward and six reverse error terms yields 12 error terms for two-port devices

10 – Network Analysis
H7215A#101 v2.0
Agilent Technologies
Page 47

As a review, the systematic errors that affect reflection and transmission measurements are shown again. For a 1 port reflection measurement, directivity, source match and reflection tracking are the largest errors. For measurement of transmission, source and load mismatch, crosstalk and transmission tracking errors must be dealt with. That is a total of 6 errors for measurements in the forward direction. When the stimulus is switched to drive the DUT in the reverse direction these errors take on different values. We must account for all 12 of these unique error vectors at each stimulus frequency.

This is why we often refer to two-port calibration as twelve-term error correction.

Characteristics of Perfect System

- **Infinite Isolation**
  - **Crosstalk**
  - **Directivity**
- **No impedance mismatches**
- **Flat frequency response**

10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 48

With this knowledge of the sources of error in our measurement, we can now define the characteristics of a "perfect" system. A perfect system is one in which these errors are not present.

Accuracy Enhancement

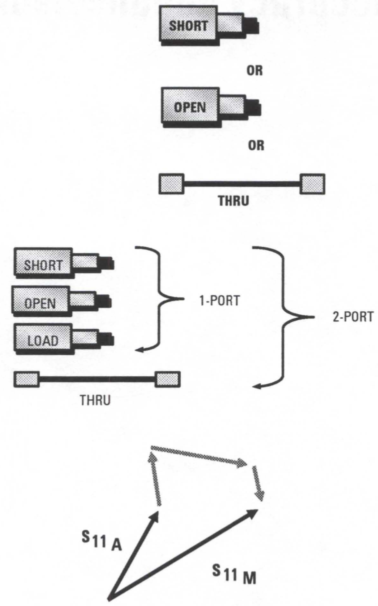
10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 49

Types of Error Correction

- **Two main types of error correction:**
  - **Response (normalization)**
    - **simple to perform**
    - **only corrects for tracking errors**
    - **stores reference trace in memory,**
    - **then does data divided by memory**
  - **Vector**
    - **requires more standards**
    - **requires an analyzer that can measure phase**
    - **accounts for all major sources of systematic error**



10 – Network Analysis
H7215A#101 v2.0

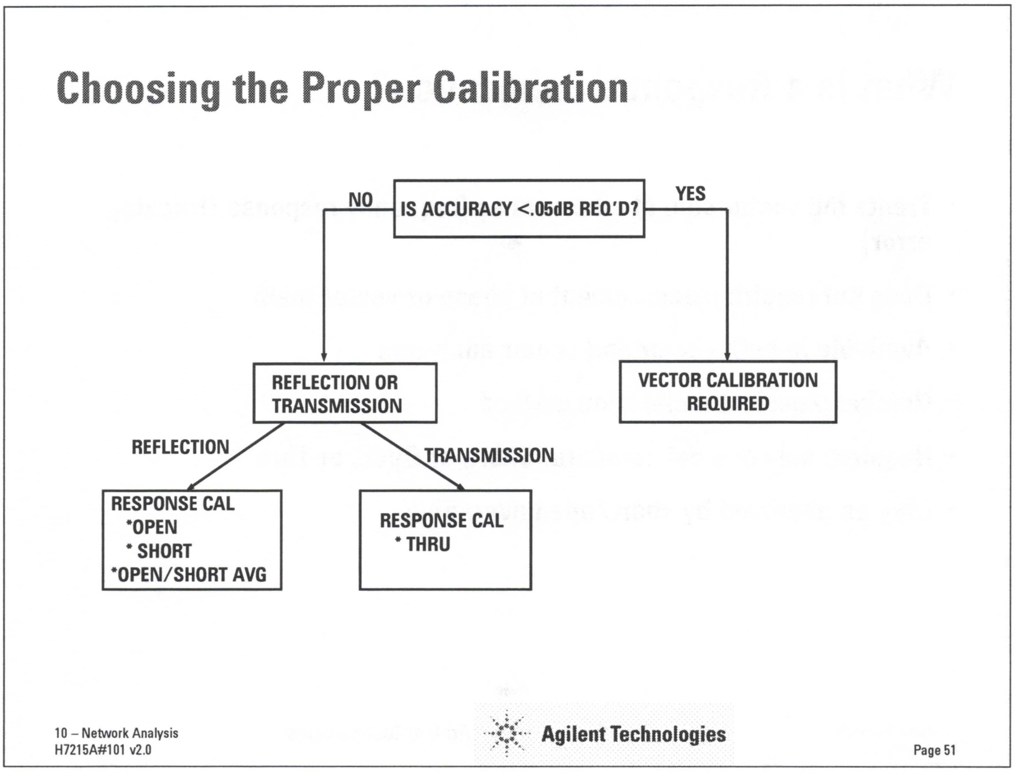
Agilent Technologies

Page 50

The two main types of error correction that can be done are response (normalization) corrections and vector corrections. Response calibration is simple to perform, but assumes that all errors are due to frequency response (tracking) and does not deal with each of the 12 systematic errors. Response calibration is a simple normalizing process where a reference trace is stored in memory, and subsequent measurement data is divided by this memory trace. Scalar analyzers utilize response calibration techniques.

Vector-error correction requires an analyzer that can measure both magnitude and phase data. It also requires measurements of more calibration standards. Vector-error correction can account for all major sources of systematic error and can give very accurate measurements.

Note that a response calibration can be performed on a vector network analyzer, in which case we store a complex (vector) reference trace in memory, so that we can display normalized magnitude or phase data. This is not the same as vector-error correction however (and not as accurate), because we are not measuring and removing the individual systematic errors, all of which are complex or vector quantities.



Modern network analyzers offer lots of calibration options . These choices afford the operator the flexibility to trade time and complexity of calibration for accuracy. When setting up for a measurement, the operator is required to consider their measurement needs and goals. If good accuracy is not a requirement, then we might choose between no calibration at all or a response calibration.

Selecting a response calibration also requires selecting either reflection or transmission measurement. One of the shortcomings of a response calibration besides reduced accuracy is that it corrects for only one measurement, reflection or transmission. For reflection, the trace may be normalized to either a short (preferred) or an open. For transmission, a thru is used.

What is a Response Calibration?

- Treats the vector sum of all errors as frequency response (tracking error)
- Does not require measurement of phase or vector math
- Available in both vector and scalar analyzers
- Quickest/easiest calibration method
- Requires only one cal standard: Short, or Open, or Thru
- May be improved by short/open averaging

10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 52

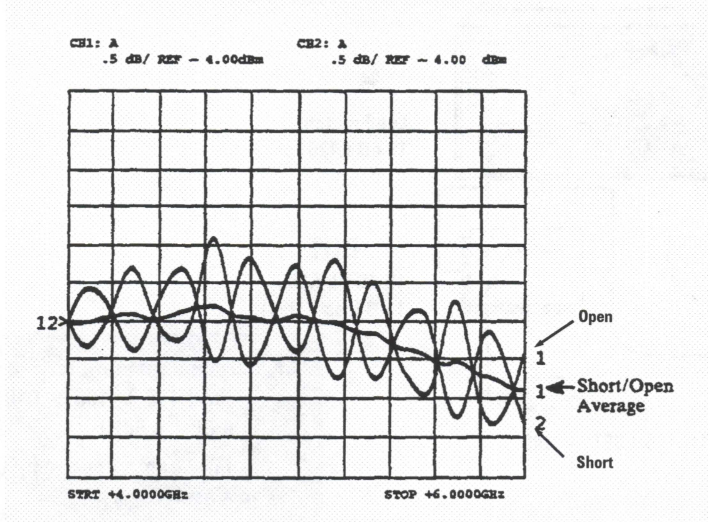
The response calibration model assumes that any deviations from the ideal measurement standard is due to frequency response error. Through simple trace normalization math, the analyzer divides the calibration trace into the measurement to produce a flat trace.

Although the response calibration does not manage all of the common systematic errors and remove them with vector math like the vector calibrations do, it still provides a quick and easy way to improve upon the uncalibrated state of the analyzer.

NOTE: Since only one error term is being corrected, only one calibration standard is required: Short (recommended) or open for reflection measurements or a thru standard for transmission measurements.

To enhance the error correction capabilities of scalar analyzers, an open/short averaging technique is used to average out the ripples caused by the source match error.

Open / Short Averaging Improves Response Cal on Scalar Analyzers

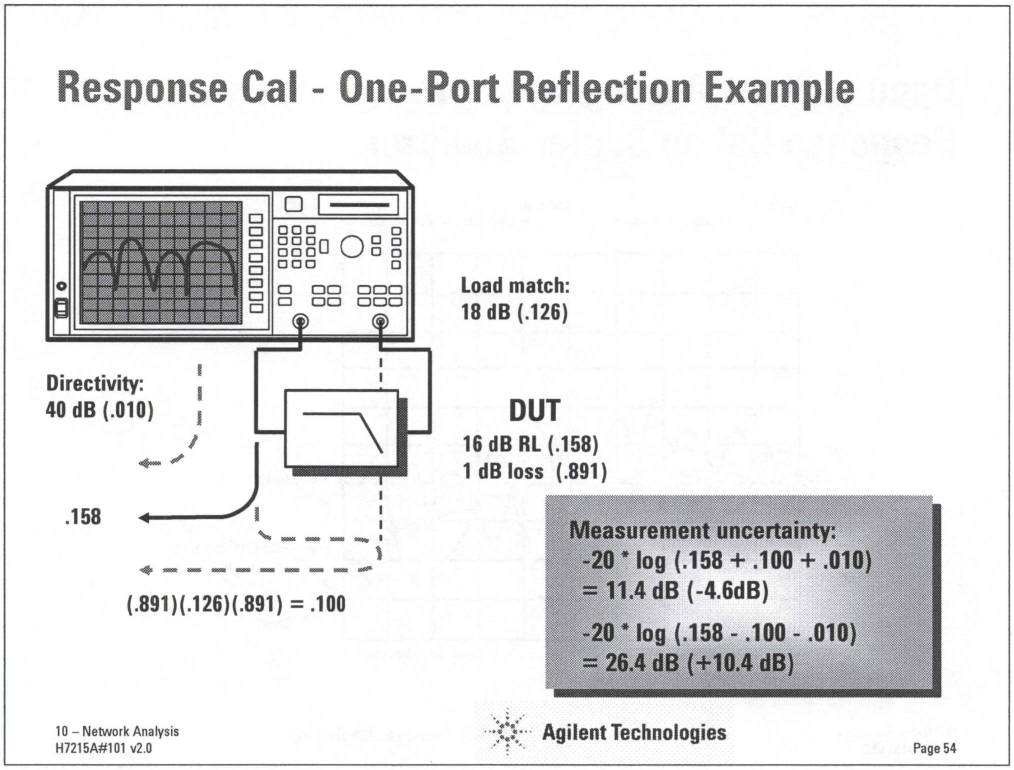


10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 53

Here is shown the effect of averaging the open and short response. The return loss of these devices should be flat at 0 dB, but because of the interaction of the reflections between the calibration standard and the imperfect match of the source, ripples are formed. These ripples are due to the vector sum of the main reflection of the standard adding in and out of phase with the re-reflections caused by imperfect match at the test port. When the other standard is applied, its reflections are shifted 180° and so are the mismatch ripples. The average of the two responses yield a net response that is closer to the actual response of either standard.



Here is an example of how much measurement uncertainty we might encounter when measuring the input match of a filter after a response calibration with a short for reflection measurement. In this example, our filter has a return loss of 16 dB, and 1 dB of insertion loss. If the raw load match of the network analyzer is specified to be 18 dB. The reflection from the test port connected to the filter's output is attenuated by twice the filter loss, which is only 2 dB total in this case. This value is not adequate to sufficiently suppress the effects of this error signal, which illustrates why low-loss devices are difficult to measure accurately.

To determine the measurement uncertainty of this example, it is necessary to add and subtract the undesired reflection signal (with a reflection coefficient of 0.100) with the signal reflecting from the DUT (0.158) (to be consistent with the next example, we will also include the effect of the directivity error signal). The measured return loss of the 16-dB filter may appear to be anywhere from 11.4 dB to 26.4 dB. In production testing, these errors could easily cause filters which met specification to fail, while filters that actually did not meet specification might pass. In tuning applications, filters could be mistuned as operators try to compensate for the measurement error. Because load match and directivity errors have not been subtracted, we are subject to the raw performance of the hardware.

Transmission Uncertainty

$$U_\tau = \frac{1 \pm \rho_S\rho_L}{(1 \pm \rho_S\rho_1)(1 \pm \rho_L\rho_2) \pm (\rho_S\tau_1\tau_2\rho_L)}$$

$$\begin{aligned} UdB &= 20\log(1 \pm \rho_S\rho_L) - \\ &[20\log(1 \pm \rho_S\rho_1 \pm \rho_L\rho_2 \pm \rho_S\rho_L\tau_1\tau_2 \mp \rho_S\rho_L\rho_1\rho_2)] \end{aligned}$$

In our example the is very small compared to the other terms

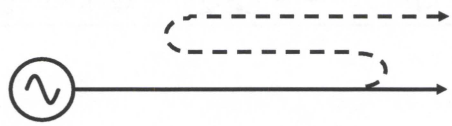
10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 55

Transmission Example Using Response Cal

- **Thru calibration (normalization) builds error into measurement due to source and load match interaction**



Calibration Uncertainty

$$= 20 \log (1 \pm \rho_s \, \rho_l)$$
$$= 20 \log (1 \pm (.200)(.126))$$
$$= \pm \, 0.22 \text{ dB}$$

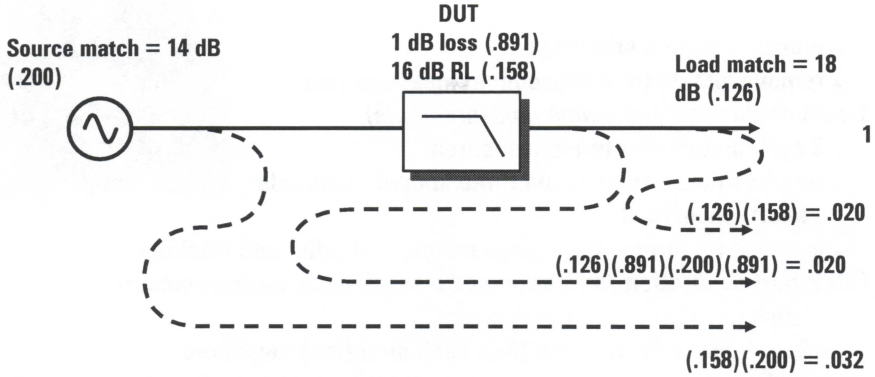
10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 56

Let's do an example transmission measurement using only response calibration. Response calibrations offer simplicity, but with some compromise in measurement accuracy. In making a filter transmission measurement using only response calibration, the first step is to make a through connection between the two test port cables (with no DUT in place). The ripple caused by the mismatch between the analyzer's source (port 1) and port 2 is calculated as $\pm$0.22 dB, and is now present in the reference data. It must be added to the uncertainty when the DUT is measured in order to compute worst-case overall measurement uncertainty.

Transmission Example (continued)



Total measurement uncertainty:
+0.60 + 0.22 = +0 .82 dB
-0.65 - 0.22 = - 0.87 dB

Measurement uncertainty
= 20 log (1 $\pm$ (.020+.020+.032))
= 20 log (1 $\pm$ .072)
= + 0.60 dB
- 0.65 dB

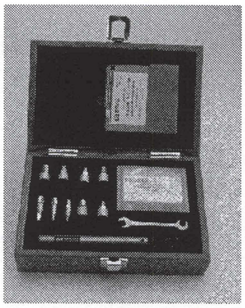
10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 57

Now let's look at the measurement uncertainty when the DUT is inserted. We will use the same loss and mismatch specifications for the DUT and analyzer as before. We have three main error signals due to reflections between the ports of the analyzer and the DUT. There are higher-order reflections present as well, but they don't add any significant error since they are small compared to the three main terms. One of the signals passes through the DUT twice, so it is attenuated by twice the loss of the DUT. The worst case is when all of the reflected error signals add together in-phase (.020 + .020 + .032 = .072). In that case, we get a measurement uncertainty of +0.60 dB, -0.65 dB. The total measurement uncertainty, which must include the 0.22 dB of error incorporated into our calibration measurement, is about $\pm$ 0.85 dB.

What is Vector-Error Correction?



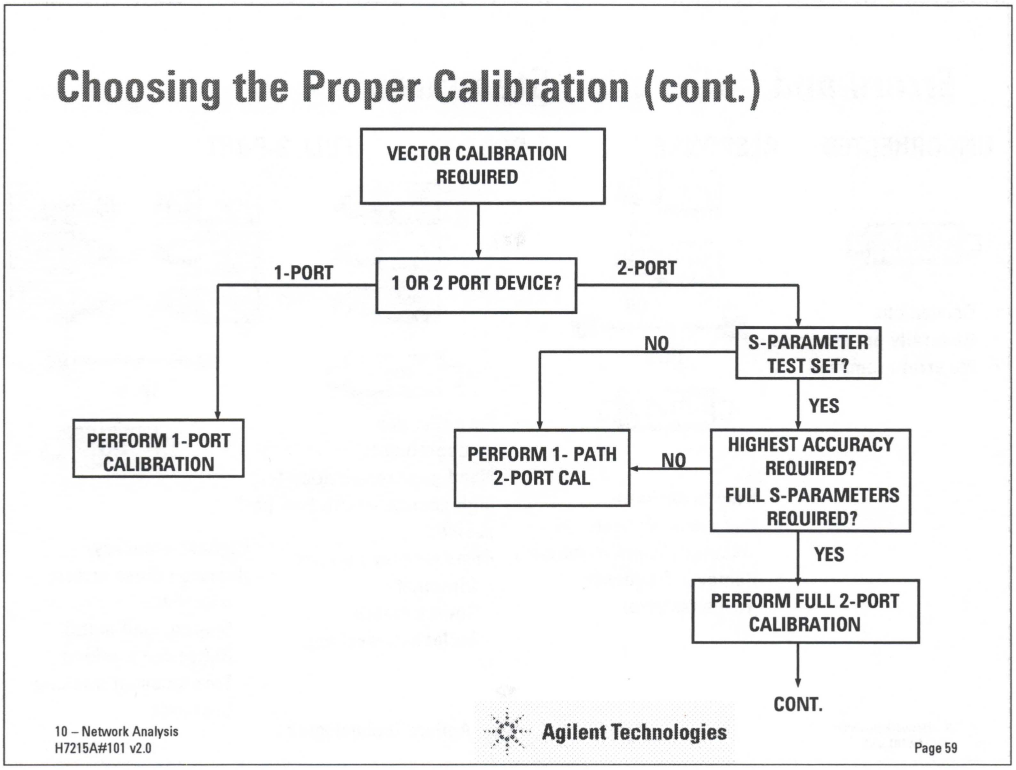
- **Process of characterizing systematic error terms**
  - **measure known standards**
  - **remove effects from subsequent measurements.**
- **1-port calibration *(reflection measurements)***
  - **3 systematic error terms measured**
  - **requires measurements on three known standards**
    - **short, open, load**
  - **corrects for directivity, source match, and reflection tracking**
- **Full 2-port calibration *(reflection and transmission measurements)***
  - **is an extension of the 1 port model**
  - **12 systematic error terms (6 in each direction) measured**
  - **requires 12 or more measurements on known standards**
    - **short, open, load, thru**
    - **thru, reflect, line**
- **Standards defined in cal kit definition file**
  - **network analyzer contains standard cal kit definitions**
- **CAL KIT DEFINITION MUST MATCH ACTUAL CAL KIT USED!**

10 – Network Analysis
H7215A#101 v2.0
Agilent Technologies
Page 58

Vector-error correction is the process of characterizing systematic error terms by measuring known calibration standards, and then removing the effects of these errors from subsequent measurements.

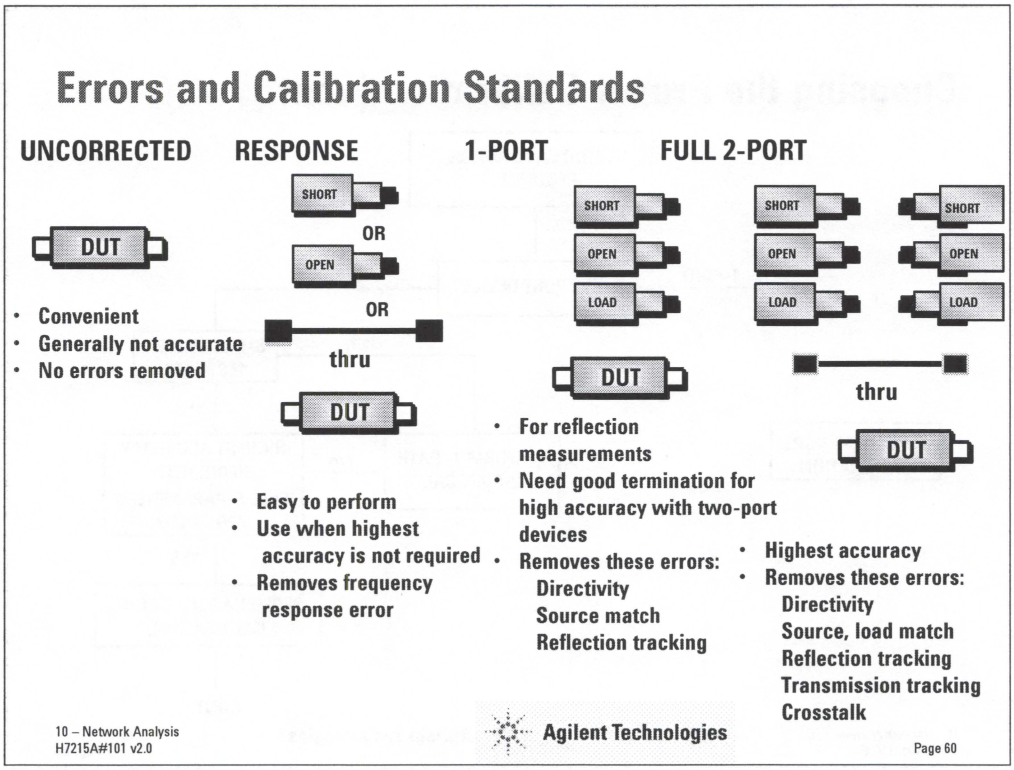
One-port calibration is used for reflection measurements and can measure and remove three systematic error terms (directivity, source match, and reflection tracking). Full two-port calibration can be used for both reflection and transmission measurements, and all twelve systematic error terms are measured and removed. Two-port calibration is an extension of the one-port model and usually requires twelve measurements on four known standards (short-open-load-thru). Some standards are measured multiple times (e.g., the thru standard is usually measured four times). The standards themselves are defined in a cal-kit definition file, which is stored in the network analyzer. In order to make accurate measurements, the cal-kit definition

MUST MATCH THE ACTUAL CALIBRATION KIT USED!



If good accuracy is required, then a vector calibration should be performed, but which kind? For a one -port device the decision is easy, use a one-port cal. This will require the use of an open, a short and a load from the cal kit.

For a two port device we will want to perform the most thorough calibration that our hardware will support. If an s-parameter test set is available then a full two-port calibration is recommended. However if a reflection/ transmission test set is all that is available then a one-path two-port calibration may be performed. This calibration model corrects for 8 error terms instead of the full 12.



Here is a summary of the calibration types available to the user their benefits and the errors that are corrected for in each.

Now let's learn how to perform a vector calibration.

Vector Calibration Procedure

- Obtain a cal kit in the DUT's connector type
- Set up stimulus frequencies and power
- Consider averaging and IF bandwidth
- Select the cal kit under the cal menu
- Select the calibration type (1-port, 2-port, etc.)
- Apply the required cal standards to the reference plane(s)
- Save the calibration
- Verify the calibration

10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 61

These are the steps to performing a vector calibration.

Before setting up the analyzer, determine what connector type is on the DUT and acquire a set of calibration standards (cal kit) in that connector type.

The stimulus frequencies and power were probably already set up during the general measurement procedure but here is a reminder that the stimulus must be set up before calibration for the error terms to remain valid. This is also a good time to consider using averaging and\or a narrower analyzer IF bandwidth. Both of these features will reduce noise in the measurement and consequently in the calibration. These features also slow the measurement throughput. Experiment with the best settings for your accuracy and efficiency requirements. Remember, averaging can be turned off after calibration to return to faster measurements.

Under the CAL menu select the cal kit that matches the set of calibration standards that you are using. Then select the type of calibration to perform; 1-port, full 2-port, etc.

Apply the cal standards to the reference planes in any order that is convenient and remember to use good connection techniques and a torque wrench if possible.

Watch the measurements as they occur and look for the measurements to agree with the standard being measured. Also look for instabilities in the traces that may indicate loose connectors or bad cables. SUSPECT EVERYTHING!

Upon completion of the process, save the calibration and the instrument state.

At this point, you should verify the calibration. There is room in this procedure for many mistakes and problems. Careful verification may be the only way to detect these problems!

Reference Plane

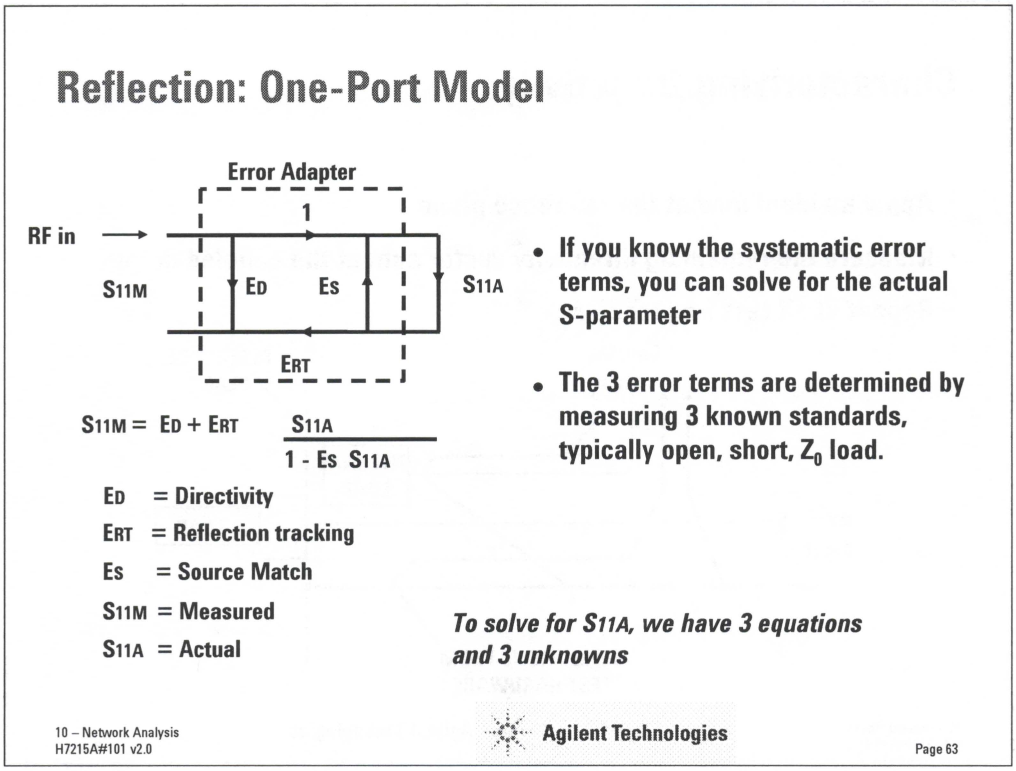
- **Defines measurement reference where calibration standards are connected (reflection or transmission)**
- **Established as close to the DUT as possible**
- **Ideal corrected characteristics :**
  - **Open circuited reflection coefficient = 1@00** 0°
  - **Short circuited reflection coefficient = 1@1800** 180°
  - **Impedance = Z0**

10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 62

The point(s) at which the test device is connected to the measurement system is defined as the reference plane(s). This is also the point where the calibration standards are connected in order to perform error correction. After error correction the reference plane will have a reflection coefficient of 1@0° when open circuited and 1@180° when short circuited. The impedance looking into the reference planes will be Z_o, the characteristic impedance of the measurement system.

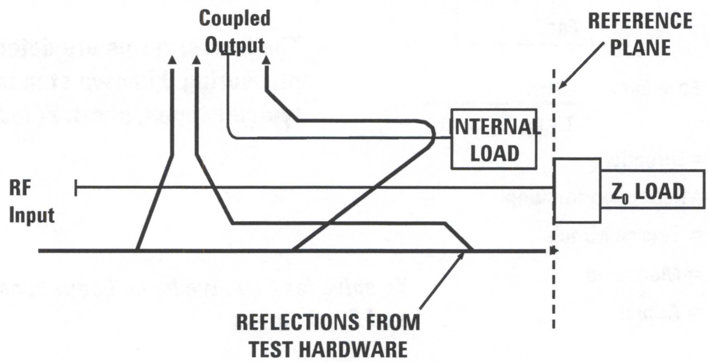


Taking the simplest case of a one-port reflection measurement, we have three systematic errors and one equation to solve. In order to do this, we must create three equations with three unknowns and solve them simultaneously. To do this, we measure three known standards, for example, a short, an open, and a Zo load. Solving the equations will yield the systematic error terms and allow us to derive the actual reflection S-parameters of the device from our measurements.

When measuring two-port devices, a one-port calibration assumes a good termination at port two of the device. If this condition is met (by connecting a load calibration standard for example), the one-port calibration is quite accurate.

Characterizing Directivity

- **Apply an ideal load at the reference plane**
- **Measure the remaining directivity vector sum at the coupled output**
- **Repeat at all test frequencies**



10 – Network Analysis
H7215A#101 v2.0

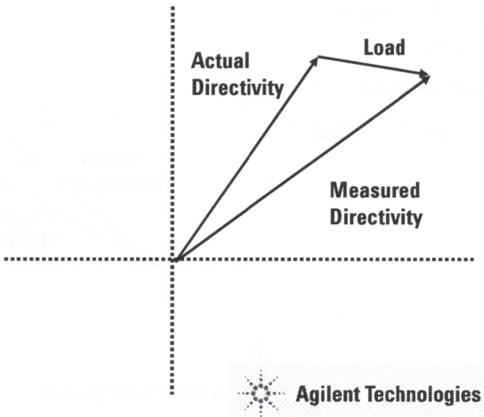
Agilent Technologies

Page 64

To determine the directivity of the network analyzer system, an ideal Z_0 load is applied to the reference plane. Under this condition, the only signals measured at the coupled port of the directional coupler (or bridge) will be due to the internal, unwanted reflections and leakages of the system. The magnitude and phase of this vector sum will be measured at each stimulus frequency and stored in an array for future correction of measurements.

Characterizing Directivity (Continued)

- Real load is not "perfect"
- Corrected directivity is the Γ of the load
- Microwave frequencies may require a sliding load



10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

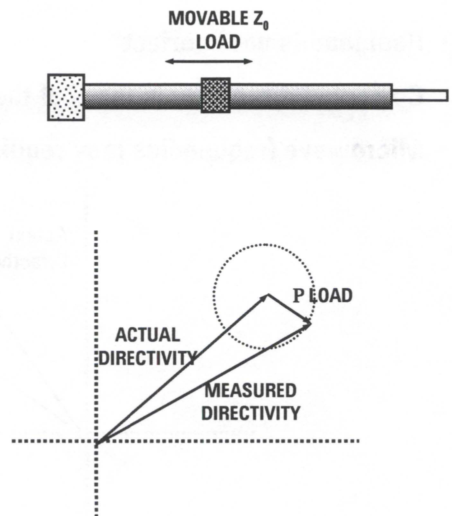
Page 65

Unfortunately no load is perfect (G=0). After correction for the directivity error is complete a residual error equal to the reflection coefficient of the load remains. This residual error represents the smallest reflection that the system can now measure. Therefore, when calibrating for directivity, we must use a load with a reflection coefficient well smaller than that of the DUT.

As the stimulus frequencies increase above UHF it becomes difficult to find a high-performance load standard. At these higher frequencies we may choose to use a sliding load for correction of directivity.

Characterizing Directivity (Continued)

- **The Sliding Load**
  - **solves for directivity at high frequencies where good fixed loads are unavailable**
  - **the phase of GLOAD is varied by moving the load element**
  - **5-12 positions are measured**
  - **calibration math solves for actual directivity**



Agilent Technologies

When measuring well matched devices at frequencies of 2 GHz and higher, the sliding load will yield better correction for the directivity error than a fixed load. It does this not because of its superior reflection coefficient but through a geometric solution for the error.

The sliding load is just a section of Z_0 transmission line with a movable Z_0 load at the end. When this load is moved, the phase of its reflection coefficient is changed causing the vector sum of the directivity error plus the G of the load to rotate. The analyzer measures this vector sum at 5 or more positions of the load and "plots" the resulting circle. By solving for the center of the circle, the actual directivity may be found. This result is then stored and used for error correction. The only residual error left over from this process is the G of the transmission line of the sliding load.

Characterizing Source Match and Reflection Tracking

1. Measure short circuit

- **S11A = 1 @~~1800~~** 180°
- **May be offset from the reference plane**

2. Measure open circuit

- **S11A = 1 @~~00~~** 0°
- **Fringing capacitance causes phase shift w/ frequency**
- **May be offset from the reference plane**

The cal kit model compensates for the offsets and phase shifts
The error correction routine solves for the error terms
from the measurement of the two cal standards

10 – Network Analysis
H7215A#101 v2.0
Agilent Technologies
Page 67

The other two errors, source match and reflection tracking are solved using two simultaneous equations and the measurements of two standards; the open and the short. The ideal reflection coefficients of these standards are well known, but due to the mechanical constraints of fabricating the standards, the analyzer may need to compensate for their non-ideal performance.

Depending on the connector type used, the short and open may not be able to be located precisely at the reference plane. Any physical offset from the reference plane will cause a phase shift from the ideal must be documented and compensated for by the analyzer during the calibration process.

The practical open will also have a tendency to radiate slightly. This phenomenon is known as fringing capacitance. This causes a phase shift vs frequency that also must be compensated for by the analyzer.

Information pertaining to these offsets is contained in the cal kit model that is supplied with each cal kit. It is therefore very important that the physical cal kit and its offset model residing inside the analyzer match in order to achieve good results.

Test-Port Connector Considerations

- **For sexed calibration kits**
  - **- open (M)**
  - **open (F)**
  - **short (M)**
  - **short (F)**
- **Refers to test-port connector (not calibration standard)**

The offsets in the calibration standards will vary with connector type and sex. For sexed connectors, the analyzer may need to differentiate the sexes of the standards being used in order to select the correct offsets to apply. When this is necessary, the calibration menu will call out male (M) or female (F). This designation is referring to the reference plane to be calibrated, not the sex of the calibration standard. For example, to calibrate a 3.5mm male connector at the end of a test cable, we would apply and open (F) and short (F) as well as a load.

Calibration Verification

- **Measure devices from a verification kit**
  - **Traceable attenuators and airlines**
- **Measure an in-house traceable verification device**
- **Re-measure calibration standards**
  - **Fixed load > 60 dB repeatability**
    - **This is not the effective directivity of the system**
  - **Short circuit < ± 0.05 Db**
    - **Phase may not be 1800 due to offset from reference plane**
  - **Open circuit < ± 0.05 dB**
    - **Be aware of capacitive and offset phase shifts**

10 – Network Analysis
H7215A#101 v2.0
Agilent Technologies
Page 69

When all of the calibration standards have applied to the reference plane and data taken from them, the analyzer will prompt the operator to save the calibration set. It may also be wise to save the entire instrument state at this time. Before measurements of the DUT begin, the calibration should be verified. It is easy to make mistakes and create errors during the calibration process. It is risky to assume that the process was trouble free and the analyzer is now measuring accurately. Verification is the only way to know for sure.

The recommended method for verification is to measure devices in which the S-parameters are known. These devices should not have been used during the calibration process and their S-parameters have been established by a metrology lab. Devices from a network analyzer verification kit are typically precision attenuators and transmission lines with air dielectrics (air lines). In-house standards may also be established.

If no independent standards are available for verification then the cal kit standards may be re-measured as a kind of "sanity check". Absolute accuracy and tracablility of the analyzer's measurements can not be verified, but some calibration mistakes and hardware problems can be found this way. Be aware that the standard offsets that were compensated for during the calibration process are not being compensated for during verification. The standards are being measured like any other DUT. For example, a 3.5mm short circuit that has an offset from the reference plane will not verify with a phase of 180 degrees. It will show a phase shift versus frequency. The open may show a phase shift from zero degrees due to its offsets and its fringing capacitance.

Vector Calibration Procedure

- Obtain a cal kit in the DUT's connector type
- Set up stimulus frequencies and power
- Consider averaging and IF bandwidth
- Select the cal kit under the cal menu
- Select the calibration type (1-port, 2-port, etc.)
- Apply the required cal standards to the reference plane(s)
- Save the calibration
- Verify the calibration

10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 70

Once again, let's review the steps to performing a vector calibration.

LAB: 1-port Calibration

Mission:

- **Perform a 1-port cal**
- **Verify the cal**
- **Compare error corrected results for S11 versus uncorrected measurements**

Need to Know:

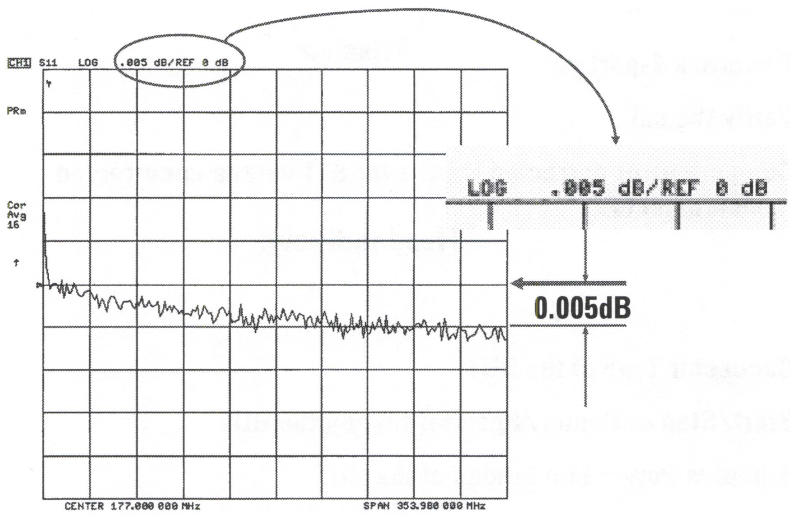
- **Connector Type of the DUT**
- **Start/Stop or Center/Span for testing the DUT**
- **Stimulus Power Limitations of the DUT**

10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 71

LAB Results: Short Verification

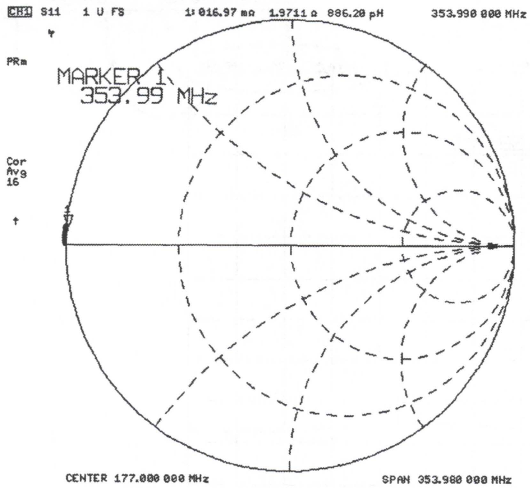


10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 72

LAB Results: Short Verification

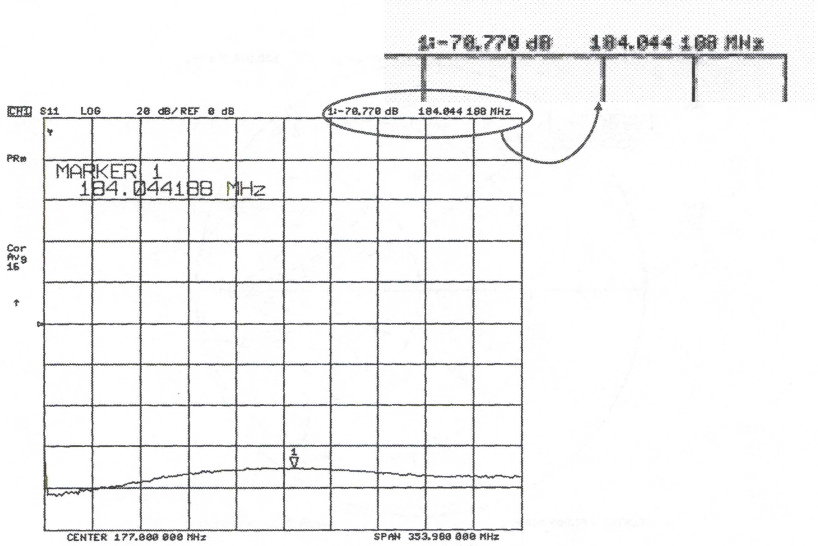


10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 73

LAB Results: Load Verification

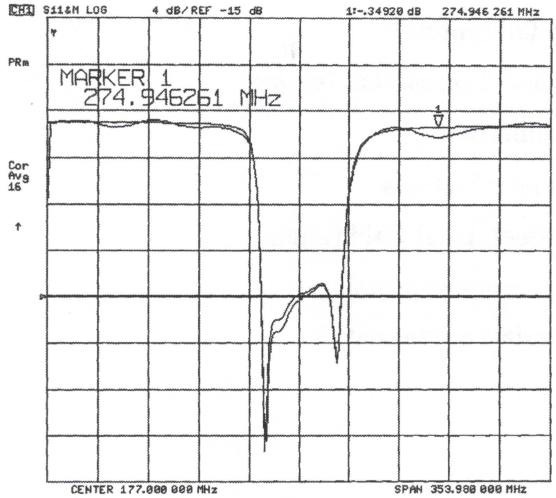


10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 74

LAB Results: S11 With & Without Calibration



10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 75

2-Port Calibrations

- **1-Path 2-Port Calibration**
  - **For Reflection/ Transmission Test sets**
- **Full 2-Port Calibration**
  - **For S-Parameter Test sets**
- **TRL (Thru, Reflect, Line) Calibration**
  - **For Highest accuracy (metrology)**
  - **For Non-coaxial environments**

10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 76

To calibrate for 2-port devices there are several choices of calibration types.

The 1-path 2-port calibration is made available for users of reflection/transmission (non-S-parameter) test sets.

The full 2-port calibration is the most commonly used method for 2 port devices.

The TRL calibration method is beyond the scope of this introduction to calibration, but it has much in common with the full 2-port calibration. TRL uses thru, reflect, and line standards instead of open, short, load and thru to solve for the systematic errors. Because these TRL standards are simple and can be manufactured in many transmission line styles, this calibration method has advantages for non-coaxial environments like stripline fixtures, on-wafer measurements and waveguide.

1-Path 2-Port Calibration

- **Equates forward/reverse terms**
- **Must reverse test device to measure S12, S22 to compute error coefficients**
- **Requires reflection cal @ port 1**
- **Requires thru standard**

10 – Network Analysis
H7215A#101 v2.0
Agilent Technologies
Page 77

The 1-path 2-port calibration method will work with non s-parameter test sets by using a simpler 8-term error model and equating some of the forward and reverse error terms. Utilizing this type of calibration requires the user to disconnect the DUT and manually turn it around in order to measure the reverse S-parameters of S_{12} and S_{22}.

The procedure steps include a 1-port reflection calibration plus the addition of a thru standard.

Full 2-Port Calibration

- An extension of the 1-port reflection cal
- Three groups of standards: Reflection, Transmission, Isolation
- Requires 2 reflection cals:
  - At both port 1 & port 2 reference planes
- Requires a THRU standard
- May require optional isolation standards
- Solves for 12 error terms (magnitude & phase @ each frequency)
- Correction for any one parameter requires measurement of all four

10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

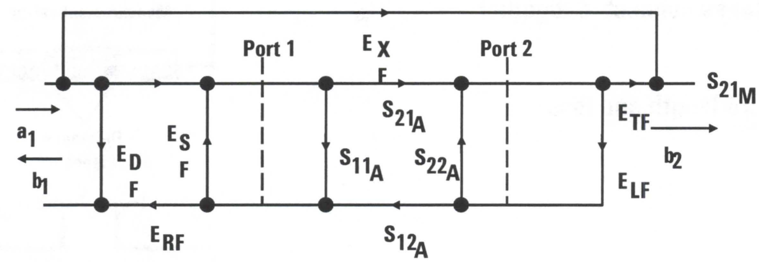
Page 78

The full 2-port cal is just an extension of the 1-port reflection calibration. The process needs 3 groups of standards to be measured; Reflection, transmission, and isolation. To satisfy the reflection group, a reflection calibration using open, short and load is performed at each of the port 1 and port 2 reference planes. The transmission group is performed by making 4 measurements of a thru standard. The isolation group acquires data from isolation standards which will be discussed later.

The result of the process is an array of 12 vector error terms for each stimulus frequency. The high accuracy of this method is gained by using the 12 error terms and the measurements of all four s-parameters to compute the corrected answer for any one s-parameter.

2-Port Error Model

Forward Direction



E_D = Directivity
E_S = Source Match
E_{RF} = Reflection Matching

E_{LF} = Load Match
E_{TF} = Transmission Tracking
E_{X_F} = Isolation

(Six more terms for reverse model)

10 – Network Analysis
H7215A#101 v2.0

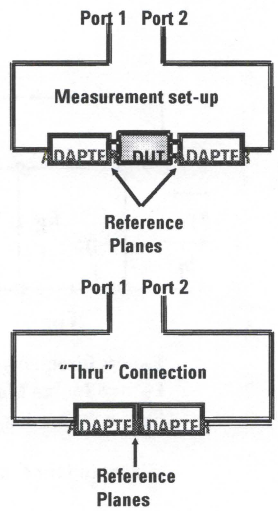
Agilent Technologies

Page 79

This is a flow diagram for the 2-port error model showing the 6 errors in the forward direction. When the stimulus power is switched to flow in the reverse direction, the 6 error terms change in value. These 12 terms are what we are trying to determine in the 2-port calibration process.

The Thru Standard

- Defined as Port 1 & Port 2 Reference planes connected together
- Zero length and loss



The difference between the 1 port cal process and the 2-port process is the addition of the thru and isolation standards. The thru standard is defined as the port 1 and port 2 reference planes connected directly together. No adapters should be added or removed for this basic thru connection. The situation may arise where the ports cannot be connected together due to the same sex connector on the ports or different types of connectors. This is called a non-insertable case and must be handled with a special procedure.

Calibrating Non-Insertable Devices

- **When doing a thru cal, normally test ports mate directly**
  - **cables can be connected directly without an adapter**
  - **result is a zero-length thru**
- **What is an insertable device?**

  - **has same type of connector, but different sex on each port**
  - **has same type of sexless connector on each port (e.g. APC-7)**
- **What is a non-insertable device?**

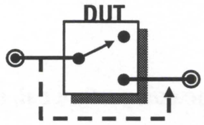
  - **one that cannot be inserted in place of a zero-length thru**
  - **has same connectors on each port (type and sex)**
  - **has different type of connector on each port (e.g., waveguide on one port, coaxial on the other)**
- **What calibration choices do I have for non-insertable devices?**
  - **Swap equal adapters**
  - **Use a characterized thru adapter (modify cal-kit definition)**
  - **Adapter removal**

10 – Network Analysis
H7215A#101 v2.0
Agilent Technologies
Page 81

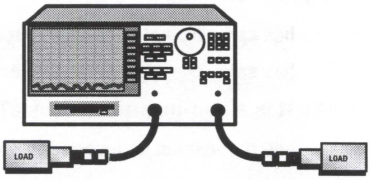
When performing a through calibration, normally the test ports mate directly. For example, two cables with the appropriate connectors can be joined without a through adapter, resulting in a zero-length through path. An insertable device is one that can be substituted for a zero-length through. This device has the same connector type on each port but of the opposite sex, or the same sexless connector on each port, either of which makes connection to the test ports quite simple. A noninsertable device is one that can not be substituted for a zero-length through. It has the same type and sex connectors on each port or a different type of connector on each port, such as N-connector at one end and SMA on the other end.

There are several calibration choices available for noninsertable devices. The first is to use a characterized through adapter (electrical length and loss specified), which requires modifying the calibration-kit definition. This will reduce (but not eliminate) source and load match errors. A high-quality through adapter (with good match) should be used since reflections from the adapter cannot be removed. The other two choices (swapping equal adapters and adapter removal) will be discussed in an advanced module.

Crosstalk (Isolation)



- **Crosstalk definition: signal leakage between ports**
- **Can be a problem with:**
  - **High-isolation devices (e.g., switch in open position)**
  - **High-dynamic range devices (some filter stopbands)**
- **Isolation calibration**
  - **Adds noise to error model (measuring noise floor of system)**
  - **Only perform if really needed (use averaging)**
  - **if crosstalk is independent of DUT match, use two terminations**
  - **if dependent on DUT match, use DUT with termination on output**



Isolation cal when crosstalk is dependent on match of DUT

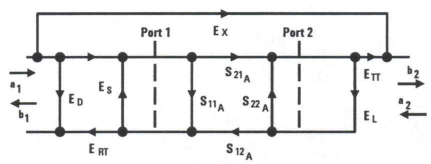
10 – Network Analysis
H7215A#101 v2.0
Agilent Technologies
Page 82

When performing a two-port calibration, the user has the option of omitting the part of the calibration that characterizes crosstalk or isolation. The definition of crosstalk is the signal leakage between test ports when no device is present. Crosstalk can be a problem with high-isolation devices (e.g., switch in open position) and high-dynamic range devices (some filter stopbands). The isolation calibration adds noise to the error model since we usually are measuring near the noise floor of the system. For this reason, one should only perform the isolation calibration if it is really needed. If the isolation portion of the calibration is done, trace averaging should be used to ensure that the system crosstalk is not obscured by noise. In some network analyzers, crosstalk can be minimized by using the alternate sweep mode instead of the chop mode (the chop mode makes measurements on both the reflection (A) and transmission (B) channels at each frequency point, whereas the alternate mode turns off the reflection receiver during the transmission measurement).

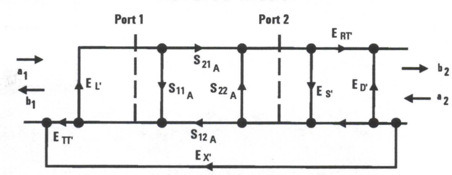
The best way to perform an isolation calibration is by placing the devices that will be measured on each test port of the network analyzer, with terminations on the other two device ports. Using this technique, the network analyzer sees the same impedance versus frequency during the isolation calibration as it will during subsequent measurements of the DUT. If this method is impractical (in test fixtures, or if only one DUT is available, for example), than placing a terminated DUT on the source port and a termination on the load port of the network analyzer is the next best alternative (the DUT and termination must be swapped for the reverse measurement). If no DUT is available or if the DUT will be tuned (which will change its port matches), then terminations should be placed on each network analyzer test port for the isolation calibration.

2-Port Error Correction

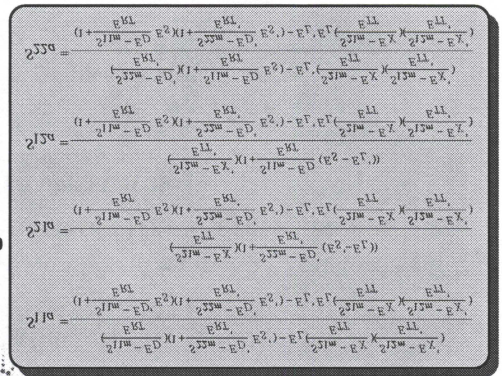
Forward model



Reverse model



E_D = Fwd Directivity	E_L = Fwd Load Match
E_S = Fwd Source Match	E_{TT} = Fwd Transmission Tracking
E_{RT} = Fwd Reflection Tracking	E_X = Fwd Isolation
$E_{D'}$ = Rev Directivity	$E_{L'}$ = Rev Load Match
$E_{S'}$ = Rev Source Match	$E_{TT'}$ = Rev Transmission Tracking
$E_{RT'}$ = Rev Reflection Tracking	$E_{X'}$ = Rev Isolation



- **Notice that each actual S-parameter is a function of all four measured S-parameters**
- **Analyzer must make forward *and* reverse sweep to update any one S-parameter**
- **Luckily, you don't need to know these equations to *use* network analyzers!!!**

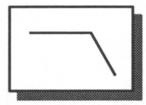
10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 83

Two-port error correction is the most accurate form of error correction since it accounts for all of the major sources of systematic error. The error model for a two-port device is shown above. Shown below are the equations to derive the actual device S-parameters from the measured S-parameters, once the systematic error terms have been characterized. Notice that each actual S-parameter is a function of all four measured S-parameters. The network analyzer must make a forward and reverse sweep to update any one S-parameter. Luckily, you don't need to know these equations to use network analyzers!!!

Calculating Measurement Uncertainty After a 2-Port Calibration



DUT
1 dB loss (.891)
16 dB RL (.158)

Reflection uncertainty

$$S_{11m} = S_{11a} \pm (E_D + S_{11a}{}^2 E_S + S_{21a} S_{12a} E_L + S_{11a}(1 - E_{RT}))$$

$$= 0.158 \pm (.0045 + 0.158^2 * .0158 + 0.891^2 * .0045 + 0.158 * .0022)$$

= 0.158 ± .0088 = 16 dB +0.53 dB, -0.44 dB (worst-case)

Transmission uncertainty

$$S_{21m} = S_{21a} \pm S_{21a}(E_I / S_{21a} + S_{11a} E_S + S_{21a} S_{12a} E_S E_L + S_{22a} E_L + (1 - E_{TT}))$$

$$= 0.891 \pm 0.891(10^{-6} / 0.891 + 0.158 * .0158 + 0.891^2 * .0158 * .0045 + 0.158 * .0045 + .003)$$

= 0.891 ± .0056 = 1 dB ±0.05 dB (worst-case)

10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 84

Here is an example of calculating measurement error after a two-port calibration has been done. Agilent provides values on network analyzer data sheets for effective directivity, source and load match, tracking, and isolation, sometimes for several different calibration kits. Here are the corrected error terms used to calculate the measurement uncertainty for this example.

Corrected error terms:

(8753D 1.3-3 GHz Type-N)

Directivity	=	47 dB
Source match	=	36 dB
Load match	=	47 dB
Refl.Tracking	=	.019 dB
Trans. tracking	=	.026 dB
Isolation	=	100 dB

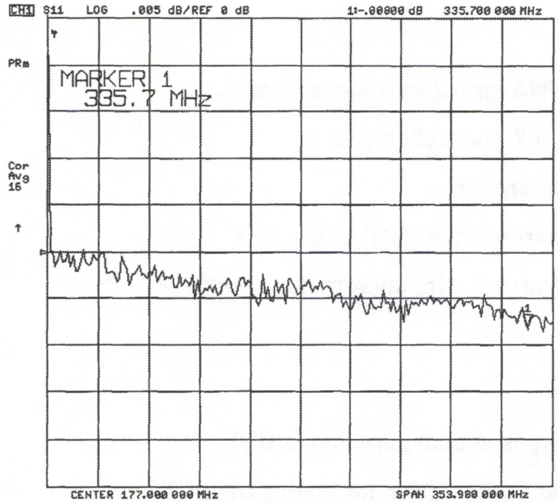
Lab: Full 2-port Calibration

- **Mission:**
  - **Set-up the VNA for a 2-port measurement**
  - **Perform a full 2-port calibration**
  - **Verify the calibration**
  - **Measure insertion loss (S21) of the DUT**
  - **Compare corrected with uncorrected results**

- **Need to Know:**
  - **Connector Types of both ports of the DUT**
  - **Start/Stop or Center/Span for testing the DUT**
  - **Stimulus Power Limitations of the DUT**

10 – Network Analysis
H7215A#101 v2.0
Agilent Technologies
Page 85

LAB Results: Short Verification

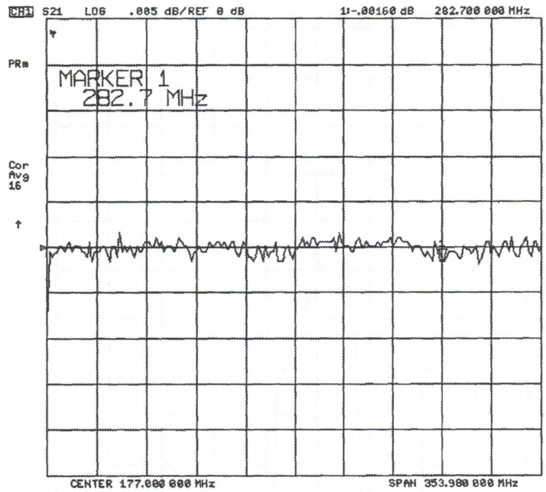


10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 86

Lab Results: Thru Verification



10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 87

Lab Results: Device S21 With/Without Cal

10 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 88

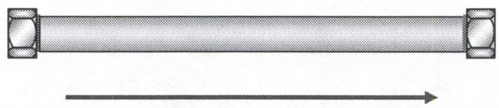
RF & Microwave Measurement Fundamentals

11 — Network Analysis

Agilent Technologies

Simple Example for Displaying Typical S21 Formats

Coaxial Cable



S_{21}

Frequency = 30 kHz to 3 GHz

11 – Network Analysis
H7215A#101 v2.0

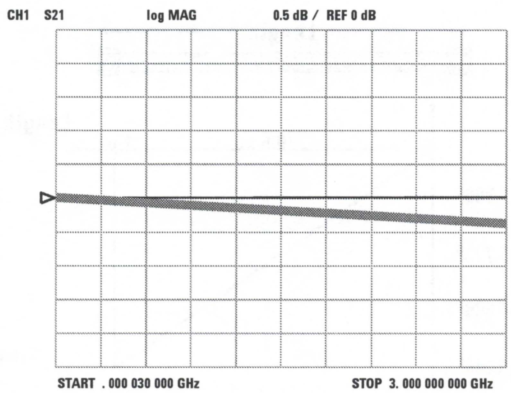
Agilent Technologies

Page 2

To examine the useful formats when making a transmission measurement (S_{21} or S_{12}), we will study the measured S_{21} of this ideal coaxial cable. At low frequencies the cable has very little insertion loss and electrical phase length but as the frequency is increased, the cable's insertion begins to increase with frequency. For this example, we will sweep the frequency from 30 kHz to 3 GHz and measure the S_{21} showing several different display formats.

Displaying S_{21} in Log MAG Format

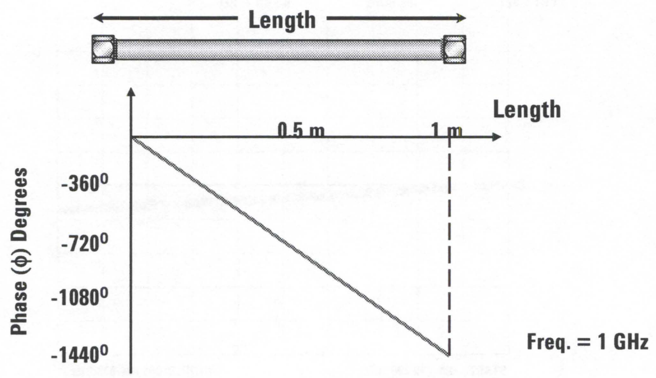
- **The S_{21} measurement with log MAG format is Insertion Gain**



The Log MAG format is the most widely used format to display the S_{21} data. This format is typically called the Insertion Loss or Gain of the DUT. This format is calculated by taking the 20 LOG_{10} of the linear transmission coefficient, τ; or $|S_{21}|$. One thing to note, the Insertion Loss of a device is typically recorded as a positive value but when the measuring transmission through a lossy device on a network analyzer, the S_{21} has a negative value. (The Insertion "Loss" implies that the negative sign is already taken into account). On the other hand, if the DUT has Gain, the gain is recorded with a positive value but the measured S_{21} would have a positive value.

Phase as a Function of Device Length

- **At a Fixed Frequency, the Phase is a function of Device Length**



11 – Network Analysis
H7215A#101 v2.0
Agilent Technologies
Page 4

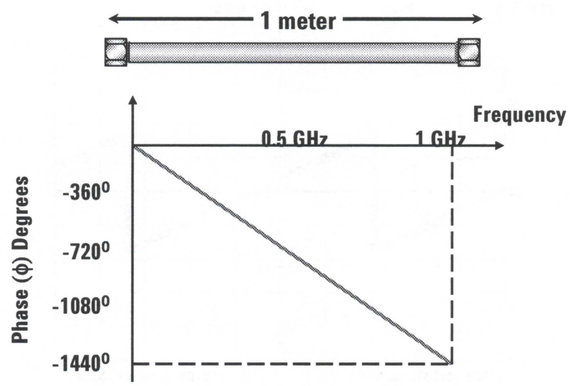
Before we examine the PHASE format on the analyzer, let's review the linear phase characteristic of a passive device. If we use a length of coaxial cable with an air dielectric and measure the phase as a function of cable length, we find that the phase angle will decrease following a certain linear slope. The cable represents a time delay to signals passing through it and the delay can be represented by some number of wavelengths. The number of wavelengths can be expressed as a phase shift measured in degrees. The formula relating device length to phase is as follows

$$\phi = -\left(\frac{360^0 \text{ f}}{\text{c}}\right)\text{L}$$

ϕ = device phase length (degs)
f = frequency (Hz)
c = speed of light (m/s)
L = cable length (m)

Phase is also a Function of Frequency

- **For a Fixed Device Length, the Phase is a function of Frequency**



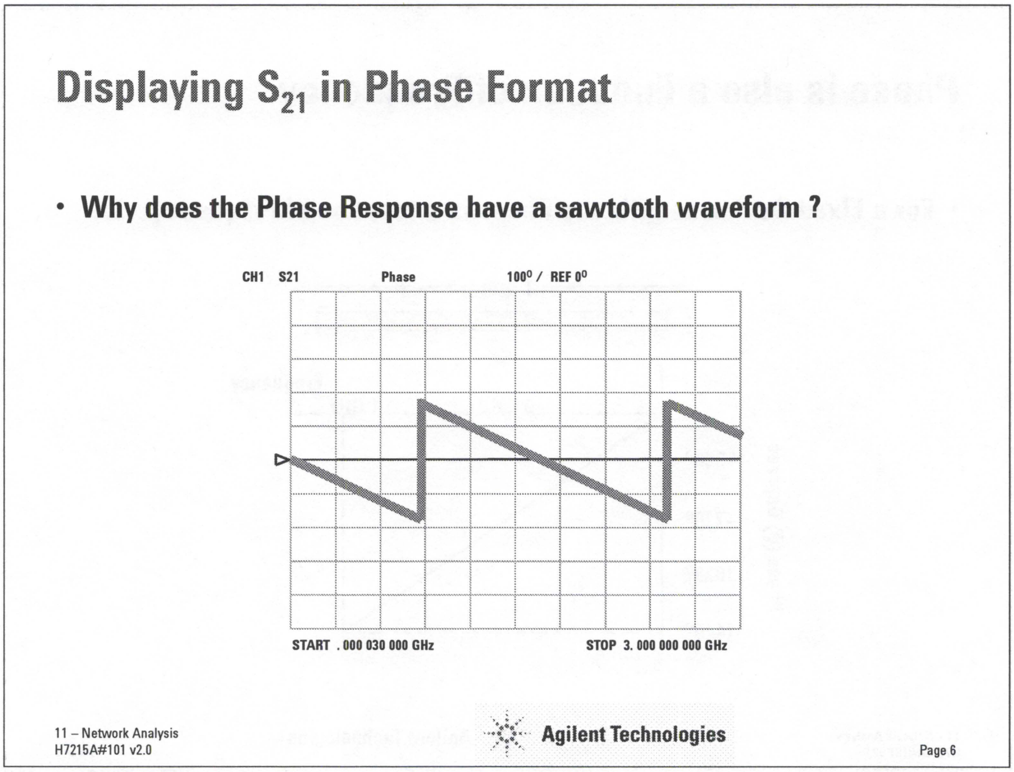
11 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 5

If we fix the length of the cable, and vary the frequency, we find the phase is also a linear function of frequency.

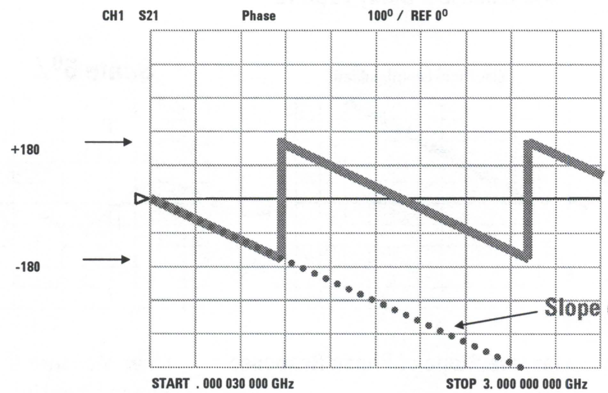
$$\phi = -\left(\frac{360^0\ L}{c}\right) f$$



Let's now look at a S_{21} measurement of a test cable using the PHASE format on the network analyzer. When measuring the phase of electrically long devices, we find a sawtooth waveform for the displayed response. Why does this curve follow a sawtooth? The reason is the analyzer is wrapping the phase between +/- 180 degrees.

The Measured Phase is Displayed from 180° to -180° (Wrapped)

- **Typical Phase Specs include Phase Length and Phase Linearity**



Why does the Phase Response have a sawtooth waveform ?

The phase response is wrapped around 0^0 covering the range $+180^0$ to -180^0

11 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 7

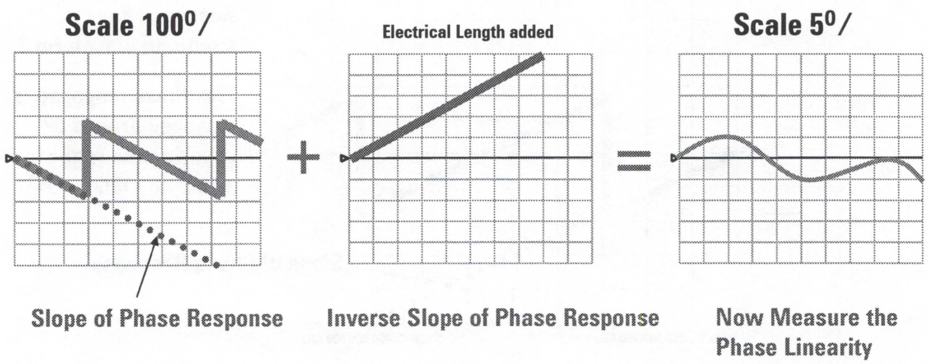
Because of the receiver architecture of the analyzer, the phase is displayed only over the range of +/- 180^0. There is no loss of information because the slope of the phase response is still maintained and the electrical delay through the device can still be determined. The analyzer will wrap the phase response when the phase is smaller than -180^0, the analyzer adds 360^0 to the measured data.

The PHASE format is typically used when comparing the phase difference between two devices or examining the phase linearity of the device across frequency. If the device has a non-linear phase response, then it may introduce signal distortion when used in a system

High Resolution Deviation From Linear Phase

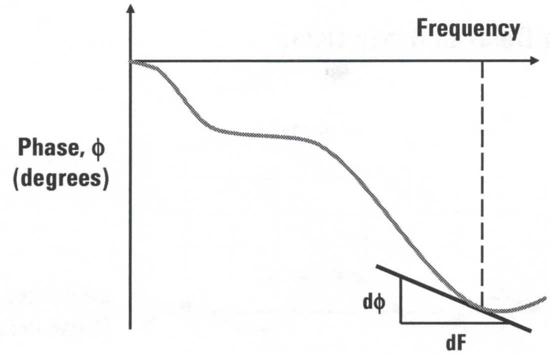
- **The Electrical Delay Feature Removes the Linear Slope from the Phase Response**

Use Electrical Delay Feature



Examining the insertion phase directly is typically not very useful. The slope of the phase response often requires a large scale factor in order to see the complete trace on the display. Also the longer the electrical length of the device, the greater the negative slope of the measured response. For cases when its important to measure the deviation from linear phase, this large scale factor often hides small deviations around the linear slope. Fortunately, the analyzer has a feature which can mathematically cancel out the linear part of the device's electrical delay, allowing the display to be re-scaled to a higher resolution. This way we can easily measure the deviation from linear phase.

Group Delay is Calculated from the Phase Response



$$\text{Group Delay} = \tau_g = \frac{-1}{360^{O}} x \frac{d\phi}{dF}$$

dF represents the Aperture setting

11 – Network Analysis
H7215A#101 v2.0

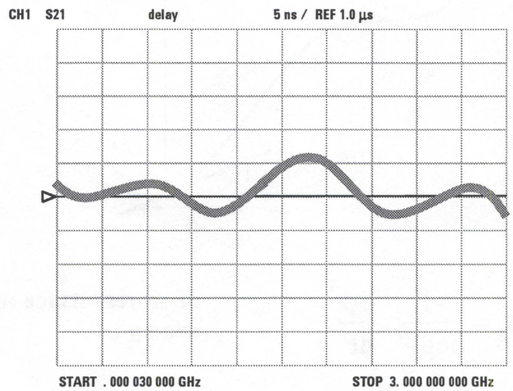
Agilent Technologies

Page 9

The group delay is calculated from the transmission phase by taking the derivative of the phase response as a function of frequency. Since the derivative is the slope of the phase response, a perfectly linear phase shift would have constant slope and therefore a constant group delay. One thing to note, when specifying or measuring delay, it is important to specify the aperture, dF, when the measurement is made. Changing the aperture can have the effect of reducing trace noise at the expense of measurement resolution. If not specified, the aperture setting should be optimized to maintain adequate delay resolution while minimizing trace noise.

Displaying S_{21} in Delay Format

- **Another name for Delay is Group Delay**



Calculated from the Phase Response

11 – Network Analysis
H7215A#101 v2.0

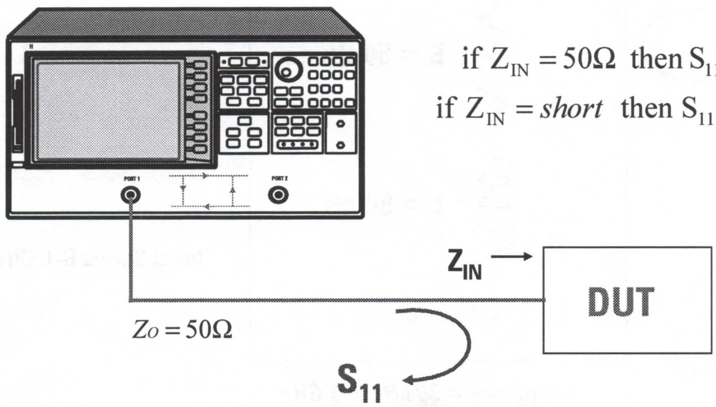
Agilent Technologies

Page 10

Another useful measurement of phase distortion is group delay. Using the DELAY format, the analyzer can now measure the group delay of the DUT. The group delay is a measure of the transit time of a signal through the device. The phase and group delay of the DUT are mathematically related. The delay is a measure of the slope of the transmission phase response. As shown earlier, if the device is electrically long, then the phase slope will be large and the resulting average transit time or average delay will also be large. The DELAY format removes the linear slope of the measured response and allows the operator an alternate way to measure phase distortion.

S_{11} Depends on the DUT's Input Impedance

- **The amount of reflected signal, S_{11}, depends on how close Z_{IN} matches 50 ohms**



$$\text{if } Z_{IN} = 50\Omega \text{ then } S_{11} = 0\angle\phi$$

$$\text{if } Z_{IN} = \textit{short} \text{ then } S_{11} = 1\angle 180^\circ$$

$$S_{11} = \frac{Z_{IN}/Z_O - 1}{Z_{IN}/Z_O + 1}$$

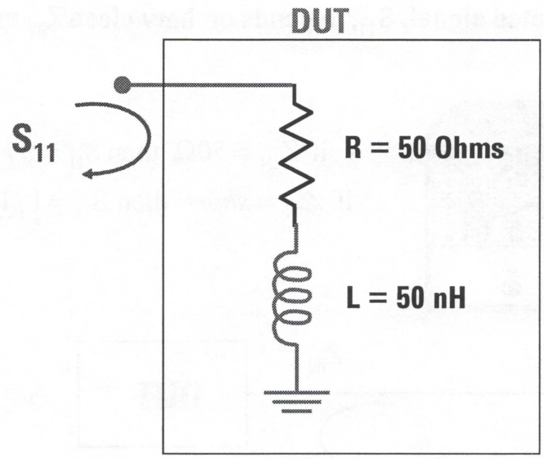
11 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 11

When measuring the reflection coefficient, Γ, or S_{11}, of a DUT, the amount of reflected signal is proportional to the input impedance of the DUT. The S_{11} is a measurement of how different the DUT's impedance is from the characteristic impedance of the measurement system, Z_o. Most high frequency test equipment has a characteristic impedance of 50 ohms. Some lower frequency analyzers can be configured with 75 ohm characteristic impedance for various CATV applications. When the DUT is a 50 ohm termination, the magnitude of S_{11}=0, with a phase angle, ϕ. If the DUT is a short circuit, all of the signal is reflected and the magnitude of S_{11}=1 (0 dB) and the phase is 180^0.

Simple Example for Displaying Typical S_{11} Formats



Ideal Series R-L Circuit

Frequency = 30 kHz to 3 GHz

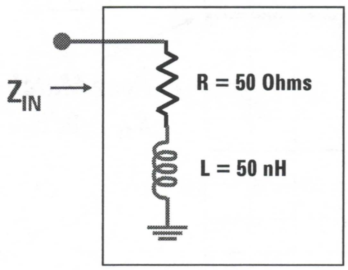
To examine the useful formats for making a reflection measurements (S_{11} or S_{22}), we will study the measured S_{11} of this simple R-L circuit. At low frequencies the inductor, L, has a very low impedance value and appears as a short to the circuit. With the inductor shorted, the DUT looks like a 50 ohm termination, so the magnitude of the reflection coefficient should be very small. At high frequencies, the inductor impedance becomes very large and the DUT begins to look like an open circuit. For this case, the reflection coefficient becomes very large. For this ideal circuit, we will sweep the frequency from 30 kHz to 3 GHz and measure the S_{11} with several different display formats.

Calculating the Input Impedance, Z_{IN}

- **For Our Example:**

$$Z_{IN} = R + jX = R + j\omega L$$

Frequency-Dependent Term



at 30 kHz

$$Z_{IN} = 50 + j0.009 \approx 50\Omega$$ ***"well-matched"***

at 3 GHz

$$Z_{IN} = 50 + j942 \approx j942\Omega$$ ***"open circuit"***

11 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

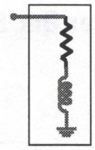
Page 13

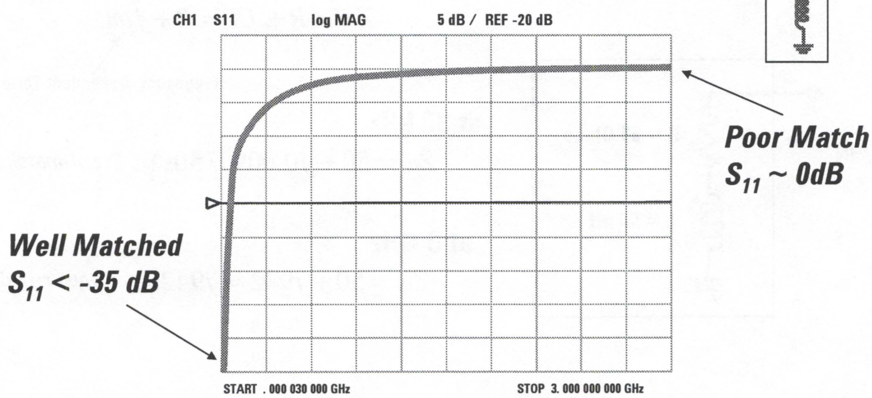
The input impedance of any DUT can be represented by a magnitude and phase or its Real and Imaginary components. The Real part of the impedance is called the Resistance, R, and the Imaginary part, X, is called the Reactance. The reactance can be inductive or capacitive depending on the device characteristics at the measured frequency. Recall the reactance of an ideal inductor is ωL and $1/\omega C$ for an ideal capacitor. The reactance is typically the frequency-dependent term of the impedance and many devices can look either inductive or capacitive depending on the measured frequency. For example, a chip capacitor has a capacitive reactance at low frequencies but as the frequency is increased above its self-resonance, the same capacitor becomes inductive as the lead and package inductance becomes larger than the specified capacitance.

In the above example of a series R-L circuit, at 30 kHz, the reactance is very small (X=0.009 ohms) and the input impedance of the circuit is approximately the circuit resistance of 50 ohms. At these frequencies, the DUT is a 50 ohm termination which is a well-matched device to the characteristic impedance of the system. As the frequency is increased, the R-L circuit begins to look like an open circuit. At 3 GHz, the impedance is approximately 942 ohms, which is a highly reflective DUT.

Displaying S_{11} in Log MAG Format

- **Another name for |S11| in dB is Return Loss**





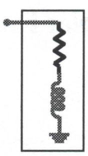
One of the most widely used display format for an S-parameter is the Log MAG format. When the Log MAG format is selected for S_{11} or S_{22} measurement, its value is typically called the Return Loss of the DUT. This format is calculated by taking the 20 LOG_{10} of the linear reflection coefficient, ρ. One thing to note, the return loss is expressed in terms of dB, and is a scalar quantity. The definition for return loss includes a negative sign so that the return loss value is always a positive number (when measuring S_{11} or S_{22} on a network analyzer with a log magnitude format, ignoring the minus sign gives the results in terms of return loss). Return loss can be thought of as the number of dB that the reflected signal is below the incident signal. Return loss varies between infinity for a Zo impedance and 0 dB for an open or short circuit.

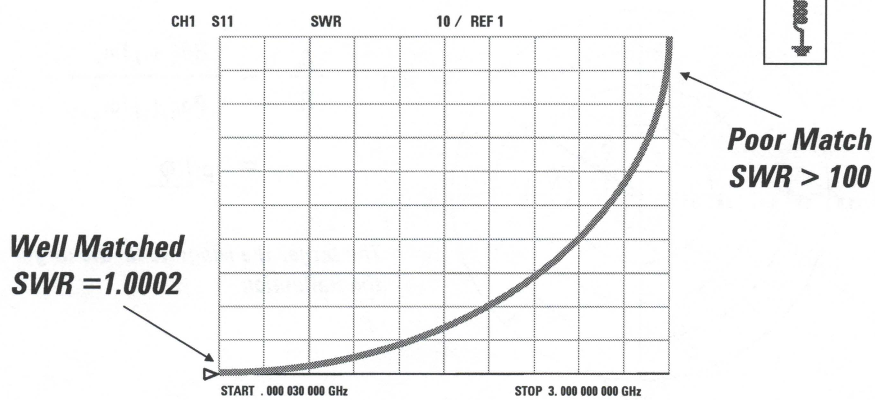
For our simple R-L circuit, the measured S_{11} on the Log MAG scale shows very good match performance at the low frequencies where the DUT's input impedance is approximately 50 ohms. The measured S_{11} is greater than -35 dB at 30 kHz. As the frequency is increased, we find the S_{11} value increasing from -35 dB to almost 0 dB. A value for S11 of 0 dB represents a complete reflection from the DUT (the reflection coefficient, ρ = 1 for a total reflection). We know that in our example, the inductive part of the circuit becomes very large and highly reflective as the frequency is increased.

$$S_{11}(dB) = 20\log_{10}(\rho)$$

Displaying S_{11} in SWR Format

- **Another name for SWR is VSWR**





11 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

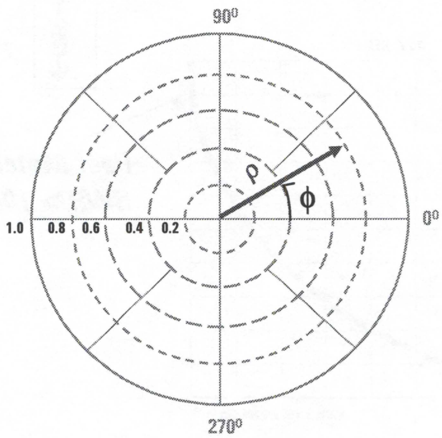
Page 15

Another very popular format for displaying the reflection from a DUT is the SWR or "standing wave ratio" format. This unit-less value represents the ratio of the maximum to minimum voltage levels that appear in the standing wave pattern created by two signals travelling in opposite directions along the same transmission line (one signal is the incident signal and the other is the reflected signal from the DUT). Many devices are specified to have a certain SWR or VSWR as determined by the manufacturer.

For our example, the above plot shows a SWR of approximately 1.0 at low frequencies where the DUT looks like a good 50 ohm termination. At higher frequencies, the inductor begins to have a very large impedance and the circuit becomes highly reflective with a SWR of 100 or greater.

Displaying S_{11} in POLAR Format

- **Polar Format shows both Magnitude and Phase on the same graph**



$$S_{11} = \frac{A}{R} = \frac{Re_A + j\,Im_A}{Re_R + j\,Im_R}$$

$$= \rho \angle \phi$$

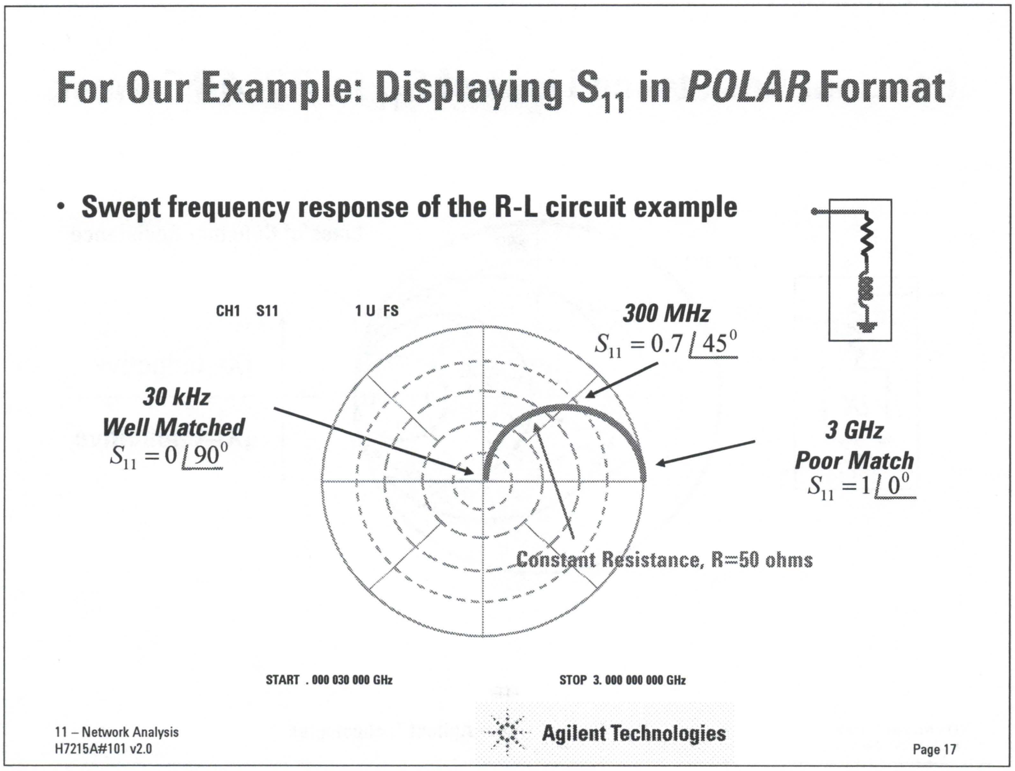
The larger the Magnitude, the larger the Reflection

11 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 16

The POLAR format is another useful S_{11} format which shows both the magnitude and phase of the reflected signal. This format displays the conversion from Real/Imaginary pairs to Magnitude/Phase values. Each point corresponds to a particular value of magnitude and phase. Quantities are read vectorally; the magnitude at any point is determined by its displacement from the center, and the phase by the angle counterclockwise from the positive x-axis. The magnitude of S_{11} covers the linear range from zero (no reflected signal) to one (complete reflection). Since there is no frequency axis, frequency information is read from markers.

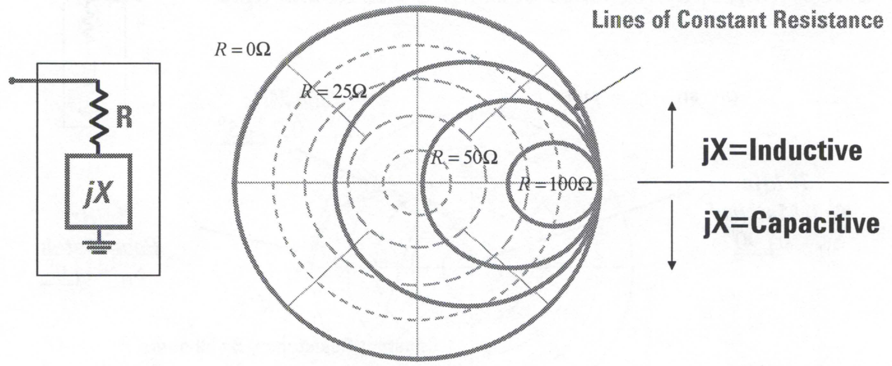


Above is the POLAR display for S_{11} using our series R-L circuit covering the frequency range of 30 kHz to 3 GHz. At low frequencies, the magnitude of S_{11} is approximately zero and the data is clustered near the center of the POLAR display. As the frequency is increased, the data is swept in an arc as the magnitude of the reflected signal is increased as the inductor's impedance becomes larger.

The phase of this circuit begins at 90^0 for 30 kHz and ends at 0^0 at 3 GHz. While the magnitude of the reflected signal can easily be seen on the SWR and Log MAG formats, the POLAR display gives some additional information which can be very useful. As an example, at high frequencies, we know this circuit becomes highly reflective (examining the SWR or Log MAG formats) but is this reflection caused by an open circuit or short circuit? Using the POLAR format, we can quickly see that the device behaves like an open circuit at high frequencies because the phase of S_{11} approaches zero. If the phase was 180^0 at high frequencies we would know that the circuit behaved as a short.

Also note, since the REAL part or resistance of this device is not changing, the arc swept by S_{11} is a line of constant resistance with R=50 ohms.

Constant Resistance Lines of S_{11} in *POLAR* Format



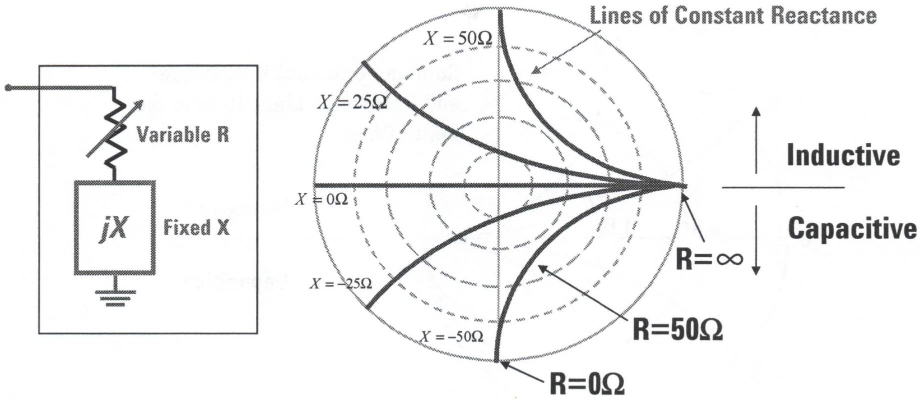
11 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 18

If we combine the S_{11} responses of the inductor and capacitor for a fixed 50 ohm resistance, we find that the frequency response will make a complete circle. For S_{11} values above the x-axis, we know the reactance is inductive and below the axis, it is capacitive. This information is very useful when designing or troubleshooting devices because we can determine not only the magnitude effects over frequency but also whether the DUT is inductive or capacitive. If we change the resistor's value, we can increase or decrease the radius of the circle as the frequency is swept. Each curve represents a line of constant resistance with a variable reactance part.

Constant Reactance Lines of S_{11} in *POLAR* Format



11 – Network Analysis
H7215A#101 v2.0

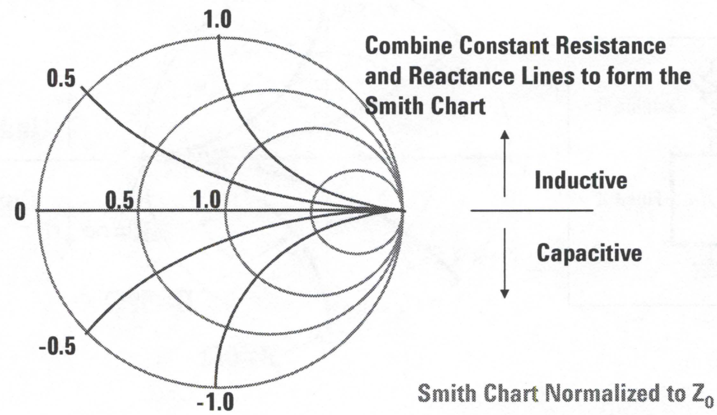
Agilent Technologies

Page 19

Above is a POLAR plot created by changing the resistance, and keeping the reactance fixed with a constant inductance or capacitance value. The above curves are functions of resistance where the open circuit point is shown to the right of the plot with a magnitude value of 1 and phase of 0^0. The short circuit has a magnitude value of 1 and a phase of 180^0 and is shown on the left side of the chart.

Smith Chart Display Format

- **Read "normalized" Z_{IN} Directly from Smith Chart**



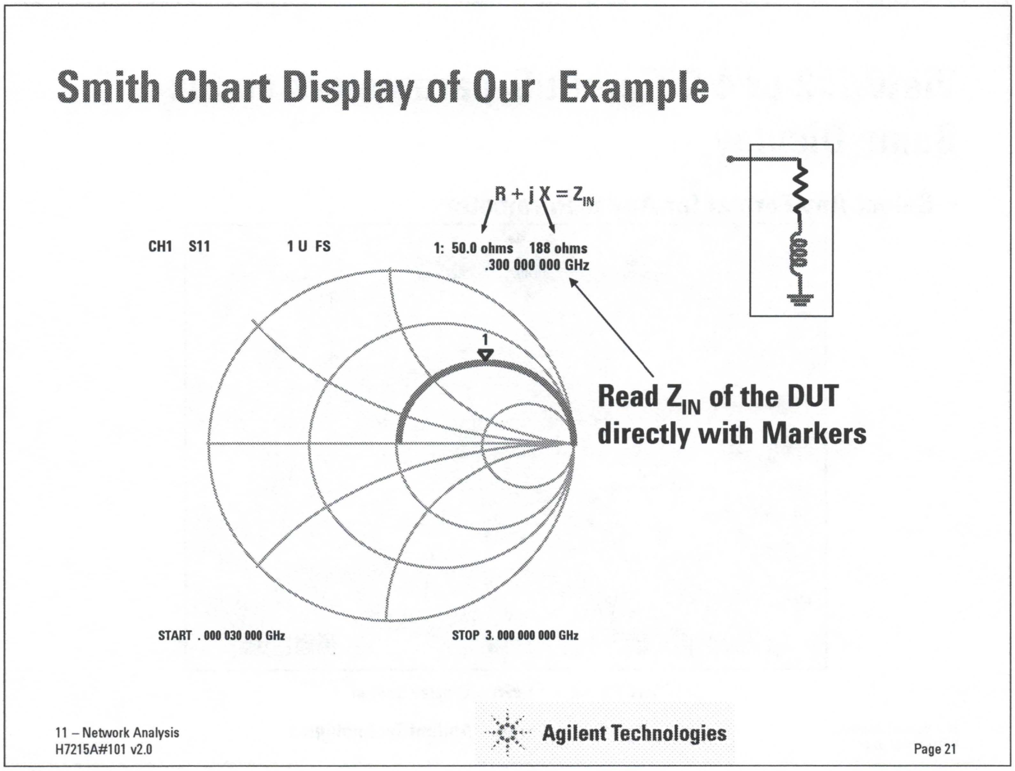
11 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 20

By taking the curves for constant resistance and constant reactance and combining them onto a common graph, we can create the popular Smith chart. This chart is typically used for reflection measurements to provide a direct readout of data in terms of impedance. Where a S_{11} measurement in a POLAR format displays the magnitude and phase of the reflected signal, the Smith chart provides the input impedance to the DUT.

The impedance displayed is normalized to the characteristic impedance of the system, typically 50 ohms. Reactance values in the upper half of the Smith chart circle have positive (inductive) reactance, and those in the lower half have negative (capacitive) reactance. The Smith chart is an excellent tool for designing and troubleshooting high frequency components which require impedance matching to the characteristic impedance of the system.



Above is the Smith chart display for our series R-L example. Using a display marker, we can find the resistance and reactance values at any specific frequency. For this case where the circuit resistance is 50 ohms, we can see that the constant resistance line of 1.0 (normalized to 50 ohms) is traced by sweeping the frequency applied to the DUT.

View 1, 2 or 4 Different S-parameters on the Same Display

- **Select Any Format for Any S-Parameter**

Two Parameter Overlay Display Shown

11 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 22

The network analyzer display can be arranged to display a single S-parameter or several on the same display. This simultaneous S-parameter viewing is very useful when tuning RF/microwave devices that change some or all their S-parameters as the device is adjusted. The display can be modified to show any S-parameter with any format. The display can also be optimized to show the S-parameter graphs in overlay or split-screen mode. Above is the analyzer display showing two S-parameters in the overlay mode.

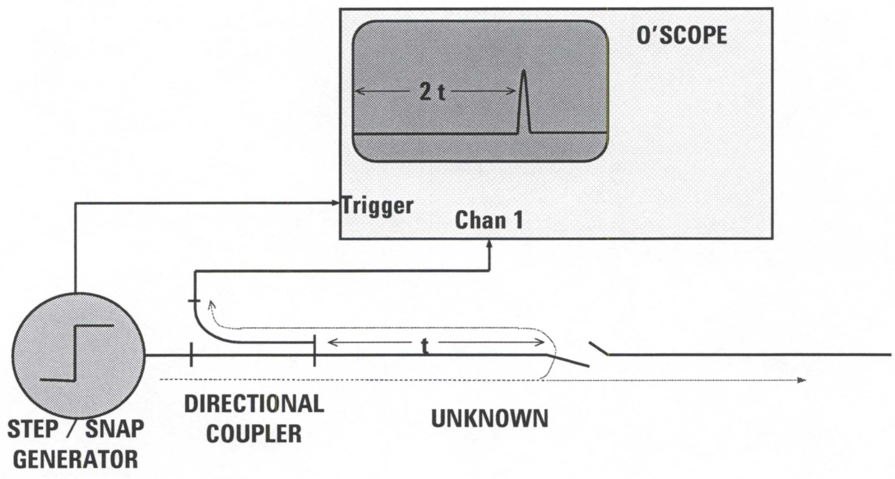
Time Domain

11 – Network Analysis
H7215A#101 v2.0

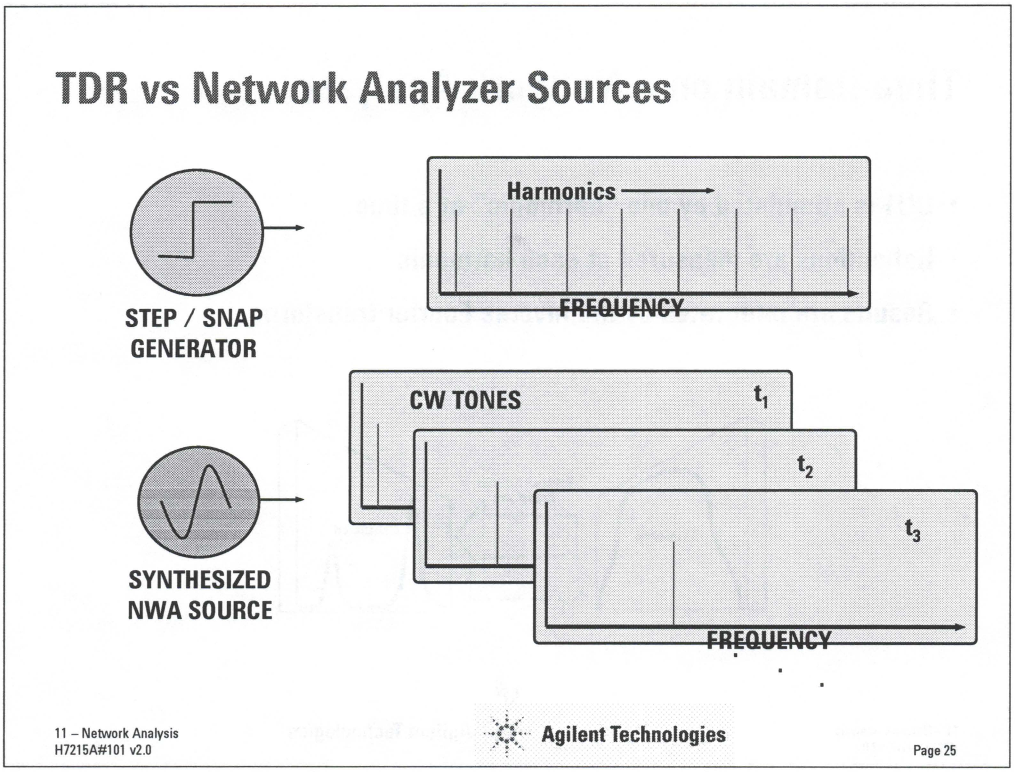
Agilent Technologies

Page 23

TDR - Time Domain Reflectometry



In the early 1960's, the technique of time domain reflectometry was introduced. This technique involves the generation of a voltage step in time that is propagated down a transmission line and coupled to an oscilloscope. If there is a difference between the characteristic impedance of the transmission line and the load impedance, a reflected voltage is generated and detected on the oscilloscope. The ratio of the measured input and reflected voltage is proportional to the impedance differences between this simple discontinuity. The measured time difference between the input and reflected voltage steps is related to the distance between the load and the source.



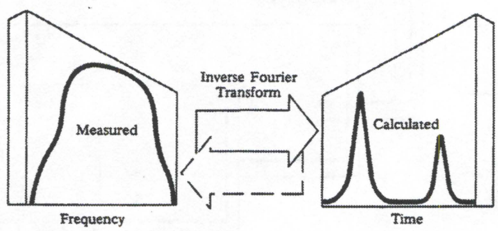
The step generator of the TDR creates a wideband signal consisting of many harmonically related spectral lines. This group of sinewave elements are generated simultaneously and the bandwidth of the group is proportional to the rise time of the step (faster risetime = wider bandwidth).

The source on the network analyzer generates sinewaves at successively higher frequencies as the "sweep" progresses from the start frequency to the stop frequency.

The two stimulus techniques achieve the same result over different time frames. All that is required for the network analyzer tp simulate a step function is to accumulate or integrate the measurement results for a number of successive frequency points.

Time Domain on a Network Analyzer

- **DUT is stimulated by one "harmonic" at a time**
- **Reflections are measured at each harmonic**
- **Results are integrated by the inverse Fourier transform**



We know that the network analyzer measures the response of an unknown at one sine frequency at a time an displays it's measurement array in the frequency domain.

Integration is the key to network analyzer's ability to emulate the TDR and produce results in the time domain. The specific integration algorithm used here is the inverse Fourier transform. The Fourier Transform is a way to decompose or separate "any" waveform into an infinite sum of sinusoids of different frequencies and amplitudes. Conversely, the inverse Fourier transform can summate the individual frequency domain elements back into a time domain response.

TDR and VNA Compared

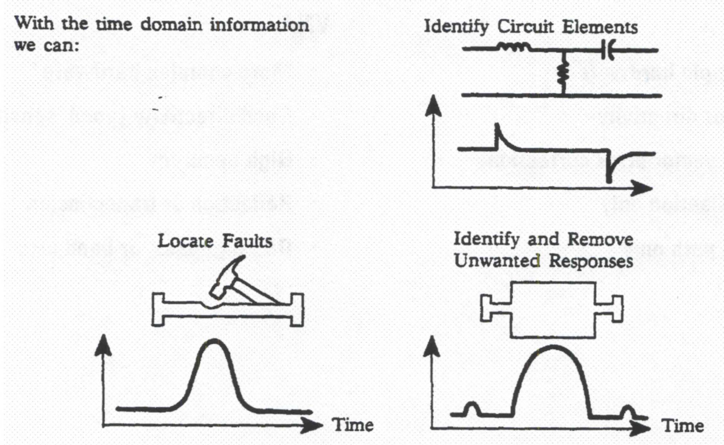
- TDR
  - Simple hardware
  - Poor directivity
  - No vector error correction
  - Reflection only
  - DC path only
- VNA
  - More complex hardware
  - Good directivity (good sensitivity)
  - High accuracy
  - Reflection or transmission
  - DC , highpass, or bandpass

TDR's still exist today and are useful to find faults in transmission lines and cables. Due to the simple hardware, the TDR makes a cost effective way to determine where to "park the backhoe and start digging". However, the TDR is not the latest in measurement technology and suffers in sensitivity, accuracy, and flexibility.

The couplers/samplers used in TDR's have poor directivity and therefore cannot see very small reflections. Without a vector receiver, vector error correction is not possible making the TDR suffer in accuracy. The TDR is also limited in application because it only makes reflection measurements of devices that pass DC, and there is no control over the frequency band of measurement.

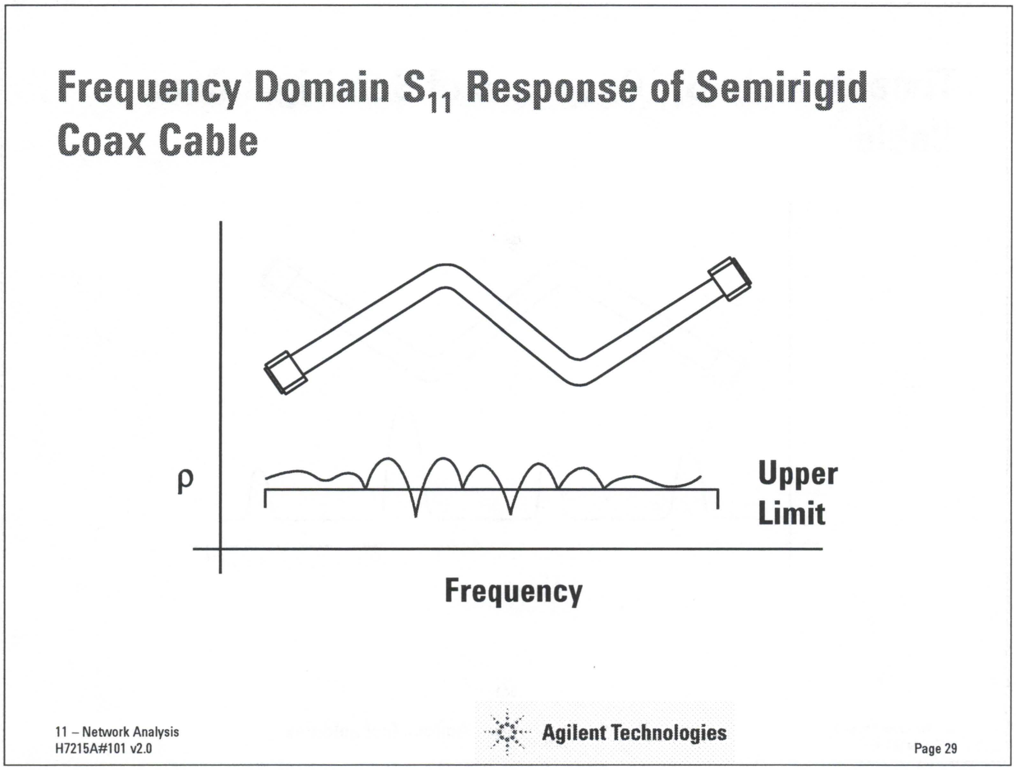
The vector network analyzer, by virtue of its more complex receiver and processing power is able to offer excellent directivity and therefore sensitivity as well as high accuracy. Measurements can be made of reflection or transmission which opens up many new applications such as the measurement of antennas, radar cross section and the analysis of surface acoustic wave (SAW) devices. The VNA's source can be easily controlled for span and bandwidth which allows the measurement of highpass, and band-limited devices.

Why the Time Domain?



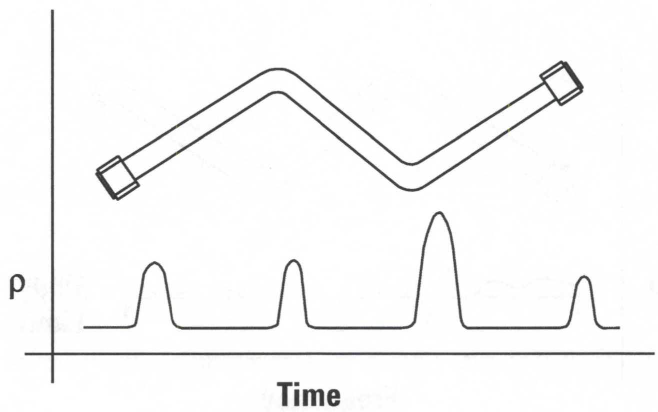
11 – Network Analysis
H7215A#101 v2.0
Agilent Technologies
Page 28

The time domain can give the operator greater insight in the the behavior of a device. When unexpected or unsatisfactory results are found in the frequency domain, further analysis in the time domain may reveal serious reflections or breaks in the transmission path, and help identify the resistive, capacitive, or inductive nature of circuit elements. An additional feature of time domain on a VNA is called gating. Gating allows the operator to isolate portions of the circuit in the time domain and view them in the frequency domain.



A good example of the application of the time domain might be the analysis of a semi-rigid coaxial cable. When reviewing the return loss of the cable, reflections in excess of the design limit are noted. Unfortunately, it is impossible to determine why this cable is our of spec. The trouble could be at the connectors, the bends or a manufacturing defect in the cable itself.

Time Domain S_{11} Response of Semirigid Coax Cable



11 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 30

After selecting the time domain and the inverse Fourier transform has been computed, a time domain trace is displayed. When compared to the physical cable, it is easy to see the reflection responses of the connectors and bends of the cable. On the vertical axis is the reflection coefficient and on the horizontal axis is time (proportional to distance to reflection). In this case, the largest reflection is coming from the second bend (left to right). Now we have an idea of where to concentrate our troubleshooting efforts.

General Time Domain Procedure

- Set-up the VNA per the general measurement procedure
  - Give additional consideration to the frequency span*
- Calibrate and Verify
- Confirm DUT operation in the frequency domain
- Turn-on the frequency to time transformation
  - Adjust parameters; bandpass/lowpass, windowing
- Select S-parameter and format (linear to start)
- Orient yourself to where things are in time
  - Spend some time "finding things" by breaking connections and causing reflections
- Set start / stop or center / span time
- Use markers and delta markers for measurement
- Consider using "time gating"

*To be covered later

Here is a generalized procedure for using time domain analysis on a vector network analyzer.

First and most importantly, the analyzer needs to be set-up, calibrated and verified to be making good measurements in the frequency domain. It is the frequency domain data from the device under test that will be sent to the inverse Fourier transform and since "garbage in = garbage out", the frequency domain data must be reliable for the time domain analysis to make sense.

Next find the time transform menu, generally under the "system" key and turn the transform "on". This will start the inverse Fourier process running with a default "impulse/bandpass" stimulus. This and other parameters may be adjusted later.

Select the S-parameter to view (typically reflection) and a format to view it in. The linear format is good for sorting out large reflections from small reflections. Try the autoscale feature.

Next, investigate ways to reconcile artifacts on the time domain trace with the physical DUT. This may be done by causing reflections to occur by loosening connections or touching the circuit with a probe.

Adjust the trace to include the responses of interest by adjusting the start/stop or center/span keys. Markers can now be used to measure the time that the stimulus traveled from the source reference plane to the measurement reference plane and the magnitude of the responses. Delta markers measure the time and amplitude difference between two responses.

Finally, the user can play "what if" by using the time domain gating feature to isolate portions of the time domain response and view that portion of the DUT in the frequency domain.

Choosing Stimulus Frequencies for Time Domain

- **What is the true frequency response of the DUT?**
- **Resolution <= 1/ freq span**
  - **Select widest valid span for best resolution**
- **Does the DUT have a DC response?**
- **Do you require R,L,C info or just magnitude?**
  - **Select "Set Frequency" feature for Low Pass Analysis**

11 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 32

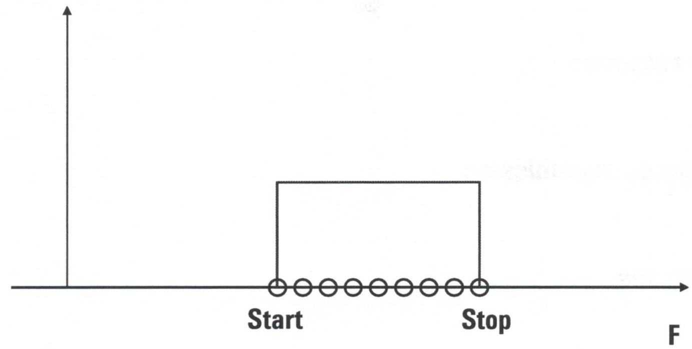
When choosing a frequency range to stimulate the DUT and then transform into the time domain some additional considerations must be made.

First, think about or even measure the true frequency response of the DUT. Even though the device may have a narrow frequency span for which it has been designed, it may have a much wider span of frequencies that it will respond to.
The reason for sweeping the device over a wider range is that the response resolution in the time domain is proportional to 1/frequency span. Resolution determines how close responses can be together and still be resolved on the time domain trace.

It is also important to know if the DUT has a D.C. response or is it a high-pass or band-pass device. If the device has a D.C. response, then it is possible to analyze it using the low pass step stimulus feature of the analyzer. Also consider whether information about the resistive, inductive or capacitive nature of the responses is of interest. If the device responds to D.C. and more information is required about the time domain responses, then select an appropriate stop frequency for the device and then press the "set frequency low pass" key under the transform menu. This will automatically choose a frequency set appropriate for low pass step analysis.

Band Pass Measurements

- **Require equally spaced points**



11 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 33

The bandpass mode of the time domain is the easiest to use. This mode simulates an impulse stimulus. To do this, all that is required is a set of stimulus frequency points over any frequency span that are evenly spaced. This is the default mode of the network analyzer. This means that any span can be used in time domain band pass.

Band Pass Mode Features

- **Works with any DUT including bandpass devices**
- **Impulse response**
- **Reflection or transmission**
- **Easiest to use**

11 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 34

The band pass time will work with any DUT over any frequency span. Remember, though that the narrower the frequency span, the less resolution in the time domain.

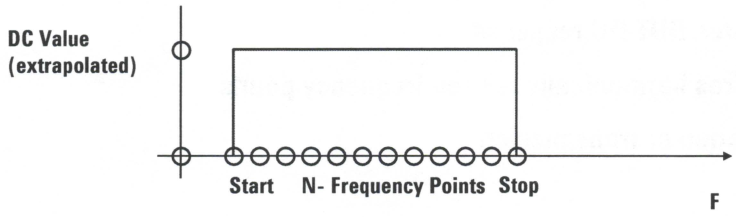
The band pass stimulus simulates an impulse. The time domain responses are therefore simple impulses. This makes the trace easier to interpret.

The band pass mode will also work in reflection or transmission. Any s-parameter measurement may be transformed into this mode.

Ease of setup and ease of interpretation makes the band pass mode easiest to use.

Low Pass Stimulus

- **Require equally spaced points**
- **Fstop = N * FStart**



11 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 35

Time domain low pass mode requires that the the frequency points be evenly spaced and that the stop frequency be equal to the start frequency times the number of frequency points (N). This makes all points harmonically related. This is a requirement of the specific form of the inverse Fourier transform being used. To insure that these criteria are met, press the "set frequency low pass" key found in the transform menu of the analyzer. This should be done prior to any calibrations.

This will set up a start frequency near D.C. and the D.C. value used in the transform will be extrapolated from the measurement at the start frequency.

This frequency set now makes it possible to simulate a step stimulus for use in the time domain low pass mode.

Low Pass Mode Features

- **Measures near-DC frequency response**
  - **Retains +- sign information of G**
  - **Identifies reflections as >< Z0 , R, L or C**
- **Requires DUT DC response**
- **Requires harmonically related frequency points**
- **Reflection or transmission**

11 – Network Analysis
H7215A#101 v2.0

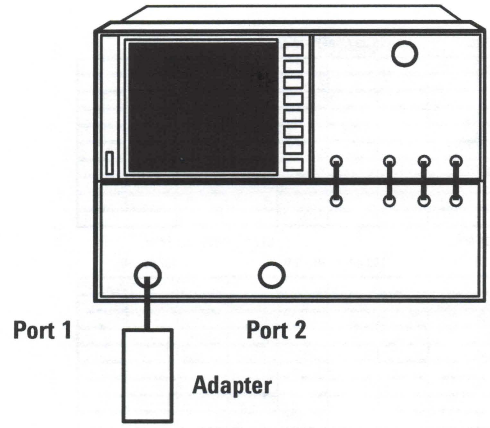
Agilent Technologies

Page 36

The low pass mode of time domain makes and estimate of the devices' D.C. response and therefore can compute the response of the DUT to a step stimulus. In a reflection measurement this means that the sign of the reflection coefficient is known and the reflection can be identified as real (resistive), inductive or capacitive. This will not only tell where and how big the reflection is but add insight as to why the reflection acts the way that it does.

This mode also works in reflection or transmission.

Open Circuit at Reference Plane



11 – Network Analysis
H7215A#101 v2.0

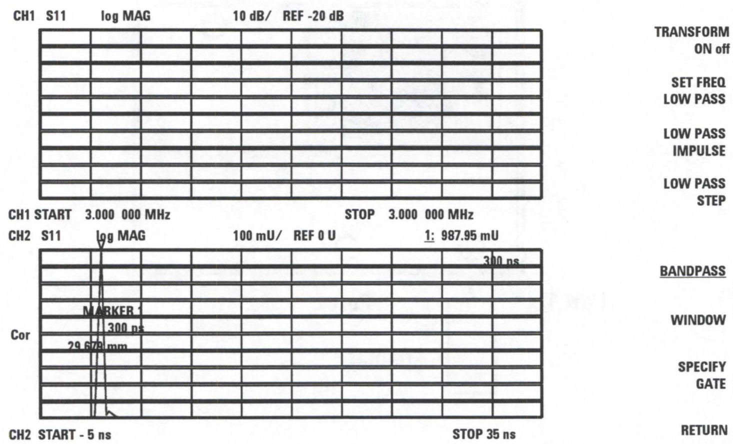
Agilent Technologies

Page 37

Now that the stimulus frequencies are set up the analyzer should be calibrated and verified to be making good measurements.

The first condition that might be seen is an open circuit at the reference plane.

Frequency Domain Response and Time Domain Band Pass Response

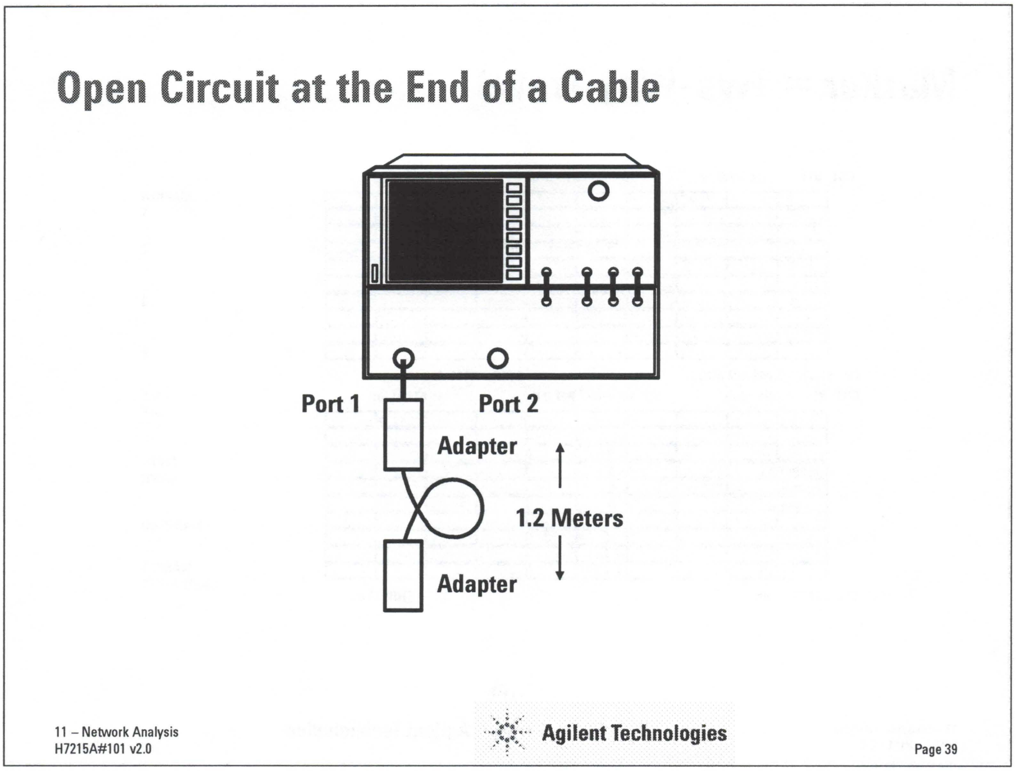


11 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

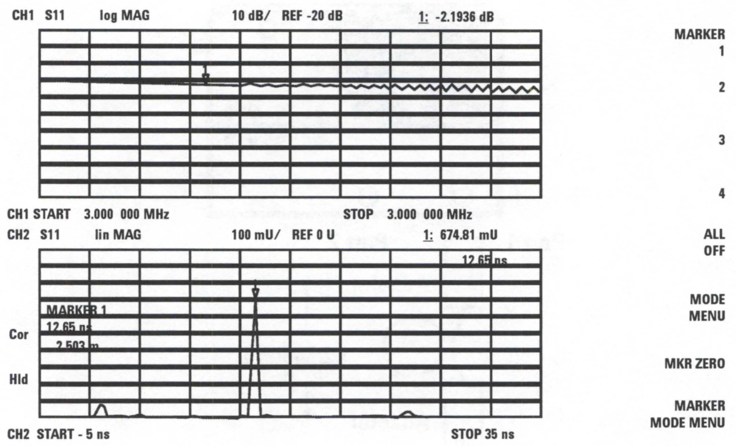
Page 38

In the frequency domain, the response of the open should measure near 0 dB return loss at all frequency points. After turning on the time domain transform and using bandpass mode, a trace similar to the bottom trace should be seen. Notice the horizontal axis s now reading in nanoseconds. The vertical axis is reading the magnitude of reflection for an S11 measurement and in this case it is a log scale. When a marker is turned on it should read the magnitude and the time from the reference plane to the reflection and back. For an open circuit at the reference plane the magnitude should be 0 dB and near 0 seconds



Moving the open to the end of cable should delay the reflection by two times the electrical length of the cable.

Marker = Two-Way Travel

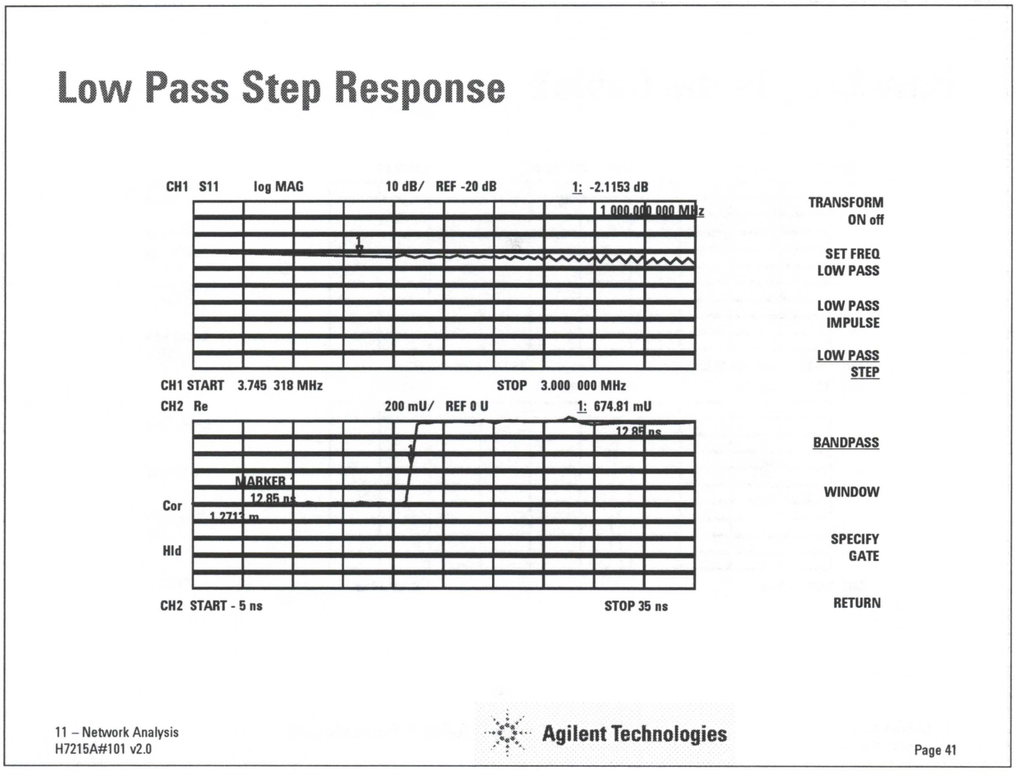


11 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

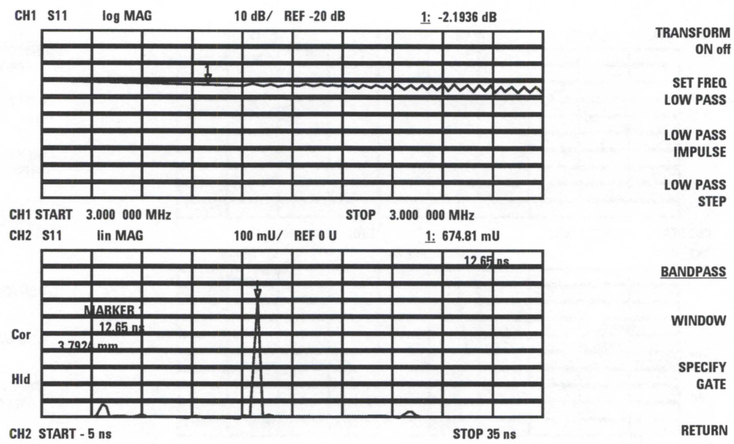
Page 40

Placing a marker on the reflection shows the round-trip time to and from the reflection. This example is viewing the reflection in the linear magnitude format. It would seem that an open circuit should have a linear reflection coefficient of one but here it reads less than that. This is due to losses and reflections that occurred before the open circuit. Both effects reduce the power incident upon the open, however the reflection measurement is made relative to the power incident upon the DUT. Therefore the open reflection reads low. This effect is called masking.



The same open circuit viewed in the low pass mode shows a step response at the plane of the open.

How Long is the Cable?



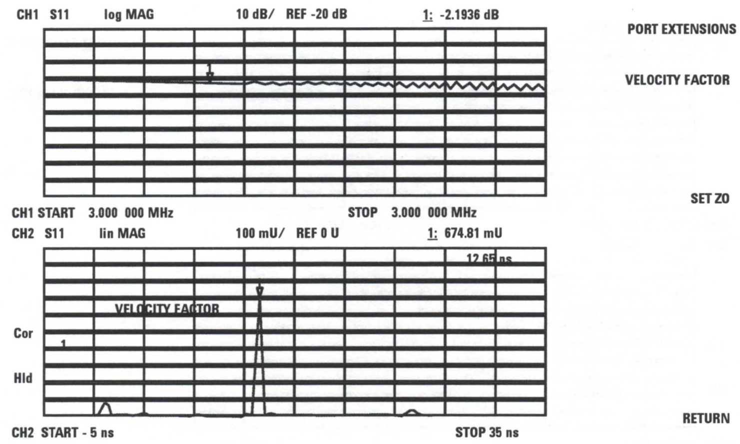
11 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 42

The marker at the plane of the open circuit reads the electrical length of the cable directly. To find the physical length of the cable, we would need to know how fast electromagnetic waves travel through this particular type of cable. To make it easier, we need only find out the speed relative to light speed through free space.

1.0 = Speed of Light in Free Space



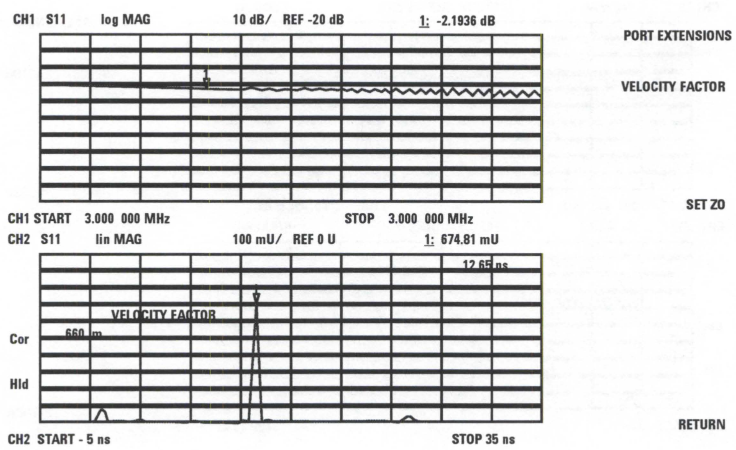
11 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 43

The ratio of speed of electromagnetic wave through a medium relative speed of light through free space is called relative velocity. The default for the network analyzer is a relative velocity of 1.0.

0.66 for Polyethlene Dielectric



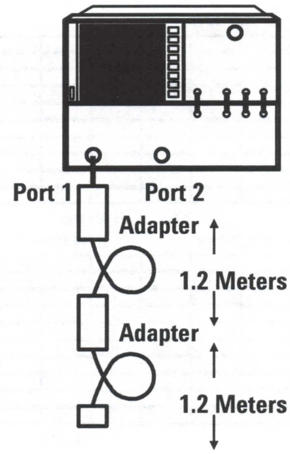
11 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 44

The relative velocity for most common dielectrics can be found in a reference book or obtained from the manufacturer of the cable. In this example, the cable has a polyethylene dielectric for which the relative velocity is 0.66. When entered into the analyzer, the marker will read both the electrical length of the cable in seconds and the approximate physical length in meters.

Open Circuit at the End of Two Cables



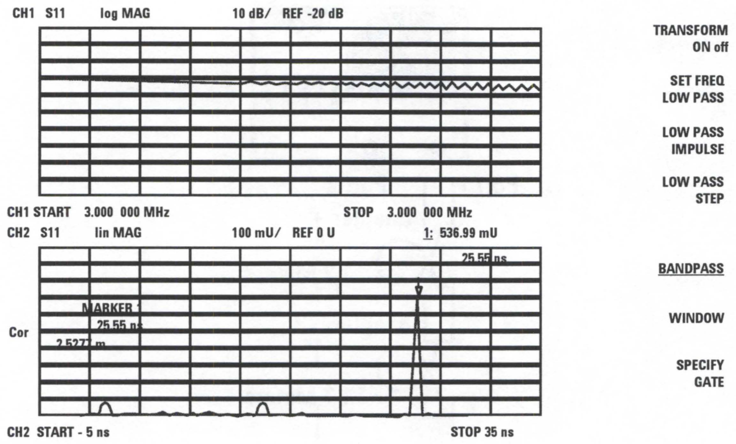
11 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 45

Adding another cable will move the open response further away from the reference plane and add another reflection from the connection of the two cables.

Open Circuit at the End of Two Cables



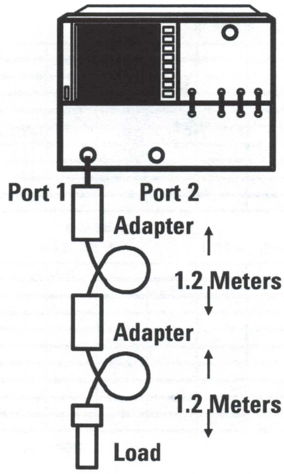
11 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 46

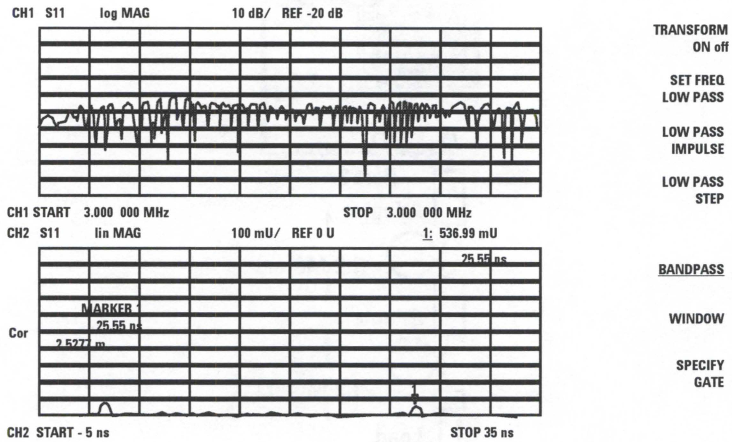
In the lower time trace, a reflection is noted at the reference plane (left), at the connection of the two cables (center), and at the open circuited end of the cable (right).

Terminated with a 50Ω Load



Adding a 50 ohm termination to the end of the cable should greatly reduce the reflection response.

Load Connected



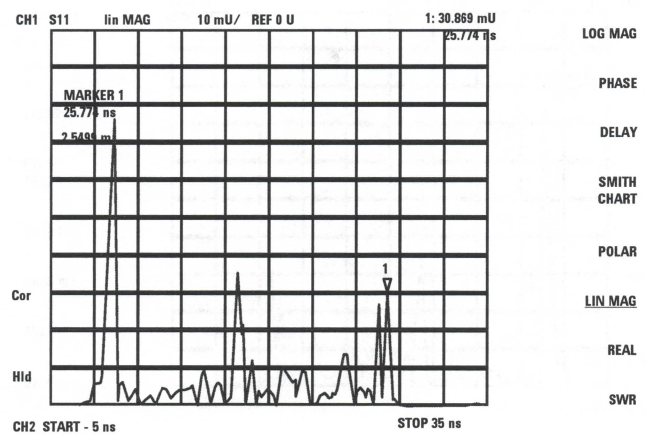
11 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 48

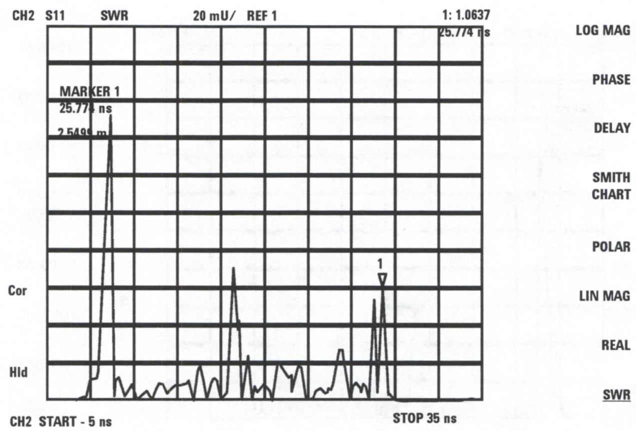
Adding the load would appear to almost extinguish the reflection from the end of the cable, but remember that the analyzer has lots of measurement range and sensitivity. Autoscaling should make these responses easier to see.

Linear Magnitude Format (Autoscaled)



After autoscaling it is easy to see the relative magnitude of the reflections. In this case it appears that the connection of the reference plane to the cable it is the most reflective. At the end of the cables, two responses are seen. The first one (left) is due to the load connector, and the second is due to the load element.

SWR Format



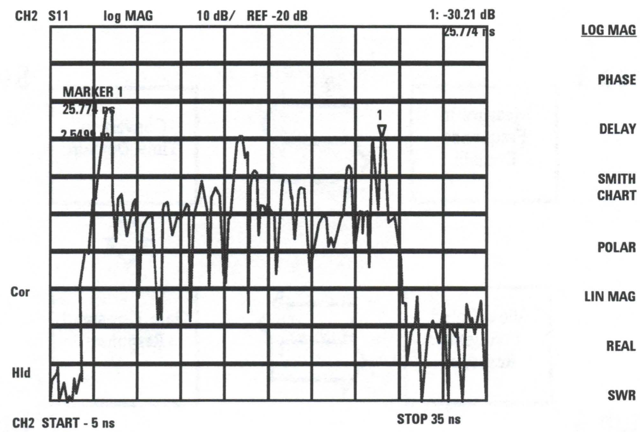
11 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 50

Although the linear display format is the easiest to interpret, other formats like SWR will also work.

Log Magnitude Format

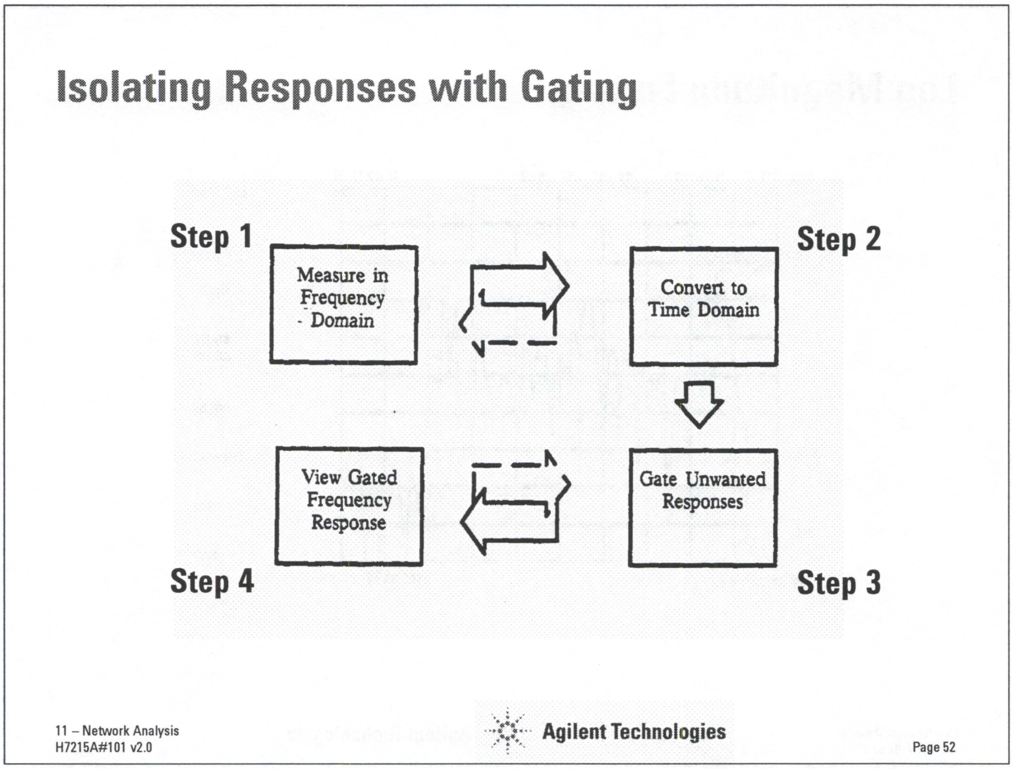


11 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 51

The log format tends to compress the peak responses and amplify the smaller reflections. Depending on the scenario, log format may help or hinder the analysis.



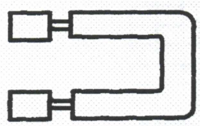
On additional powerful feature of time domain is gating. This capability allows the user to isolate portions of a transmission path in the time domain, "time pass filter" the desired responses, and then view these isolated responses transformed back into the frequency domain. This lets the user to view the s-parameters of their circuit, one element at a time or to remove unwanted responses and see if device performance improves.

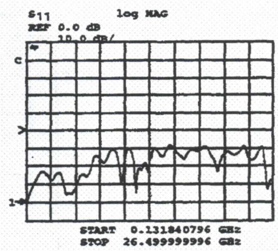
Analyzing Microstrip Launches and Bends

Step 1

Problems:

- Poor connectors
- Poor load





11 – Network Analysis
H7215A#101 v2.0

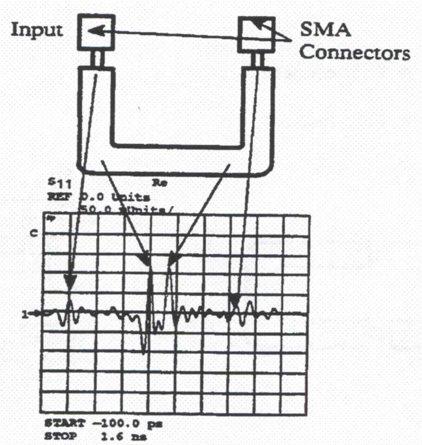
Agilent Technologies

Page 53

A common but difficult device to analyze is a simple transmission line. In this case, mismatches and losses are very small making any problems very subtle. Our example is of a microstrip transmission line with 2 bends and 2 coaxial to stripline launches at each end. When the device's return loss is viewed in the frequency domain, reflections are seen in excess of an acceptable limit. Unfortunately this composite return loss trace yields no clues as to where the reflections are coming from. A first guess would be to blame the coax launches for the reflections, but analysis will tell us more.

Convert to Time Domain

Step 2



11 – Network Analysis
H7215A#101 v2.0

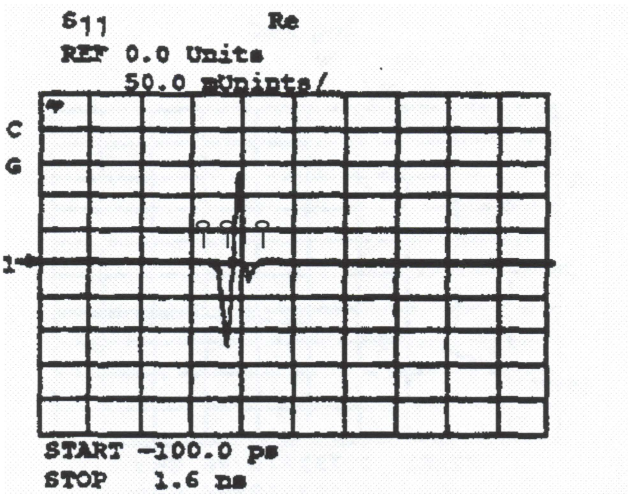
Agilent Technologies

Page 54

Viewing the time domain performance reveals the trouble to be in the stripline bends and not at the launches. Furthermore, the response of the bends show the first bend (left) to have a reactive element where the second bend appears to be resistive. Let's isolate the first bend with a gate and view its characteristics in the frequency domain.

Gate Out Unwanted Responses

Step 3



11 – Network Analysis
H7215A#101 v2.0

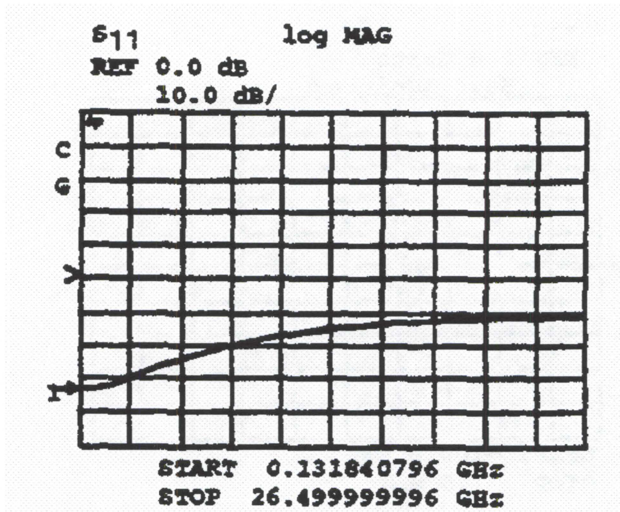
Agilent Technologies

Page 55

The markers shown on the trace describe the start, center and and stop of the gate, a time pass filter. The trace is shown after the gate has been turned on and only the first bend response remains. It is this data that will pass through the Fourier transform to the frequency domain for further analysis.

View Gated Response in the Frequency Domain

Step 4



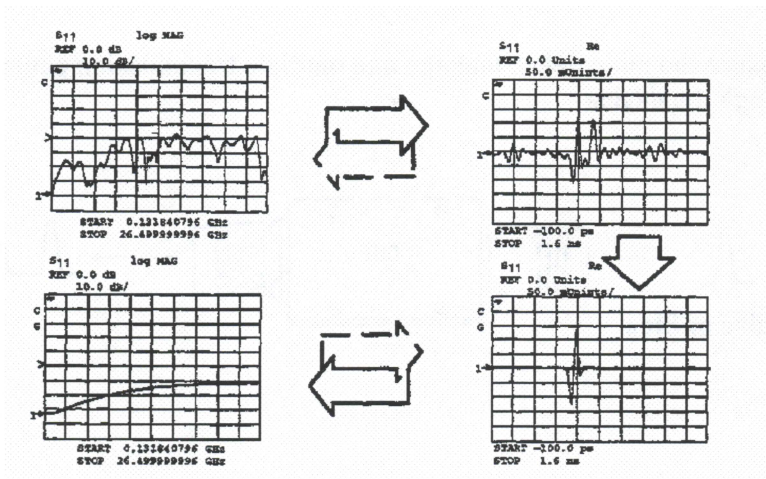
11 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 56

Here is shown the frequency response of the isolated (gated) stripline bend. This high-pass characteristic indicates that the bend has a capacitive nature and reveals more about why the bend behaves as it does. Combined with the other reflections, this characteristic was obscured.

Gating Summary



11 – Network Analysis
H7215A#101 v2.0

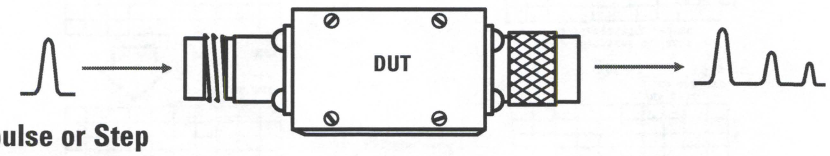
Agilent Technologies

Page 57

To summarize, the gating feature allows the user to isolate portions of a transmission path in the time domain, "time pass filter" the desired responses, and then view these isolated responses transformed back into the frequency domain.

Transmission Measurements in the Time Domain

- **Measure the main delay, leakage and multiple transmission paths through the device**



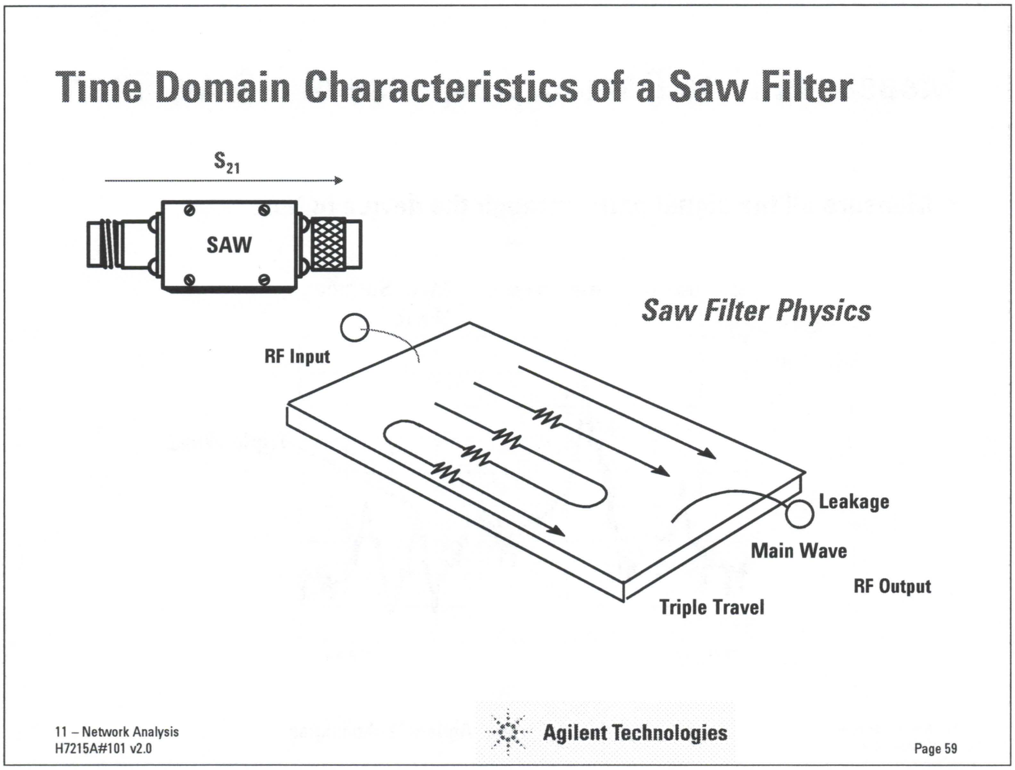
Impulse or Step

11 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 58

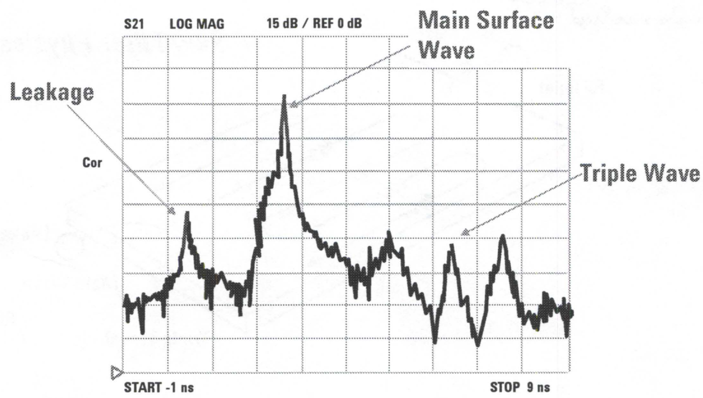
This final example of time domain serves as a reminder that the time domain function also can be useful for transmission measurements. Here we will look at a SAW (surface acoustic wave) filter. The task here is to measure the multiple transmission paths through the device and look at device performance improvements by gating the reflection responses.



A saw filter propagates waves across an acoustic substrate much like tossing a stone into a pond. These waves travel across the substrate and most of the energy is transferred off of the substrate. Some of the energy however, is reflected back to the source only to be re-reflected. This re-reflected wave shows up at the output at sometime later. Some of the energy doesn't get onto the substrate and "leaks" directly to the output. The combined output of main wave, leakage and reflected waves compromises the performance of the filter. Using the time domain transform on the forward transmission data will separate these components in time.

Measured Time Domain Response of a Saw Filter

- **Measure all the signal paths through the device in time**



11 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

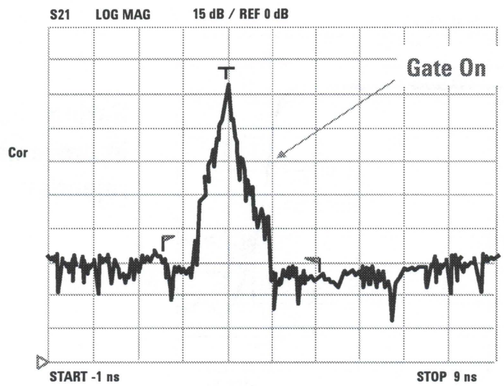
Page 60

Viewing the time domain, the shortest path in time is the leakage path directly from input to output of the device. The next response is the main surface wave and by design, it should be the largest signal at the output. Following the main wave will be one or more re-reflections, the first of which is the triple wave. We can observe here the relative magnitudes of these transmission components, but it is difficult to predict how they will combine to compromise device performance.

By gating one or more of these responses, we can observe device performance back in the frequency domain.

Gated Response of a Saw Filter

- **Gate the measured response around the main signal path**



11 – Network Analysis
H7215A#101 v2.0

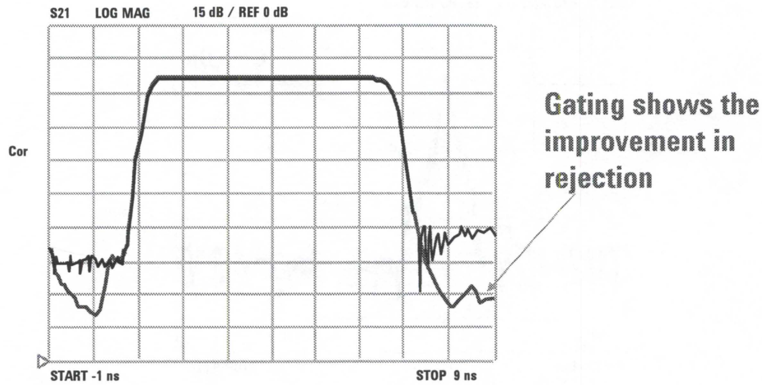
Agilent Technologies

Page 61

Here the main response has been gated. Lets view this response in the frequency domain as if there was no leakage or triple wave at the output.

Frequency Response of a Saw Filter using Gating

- **Eliminating the leakage signal improves the filter's rejection**

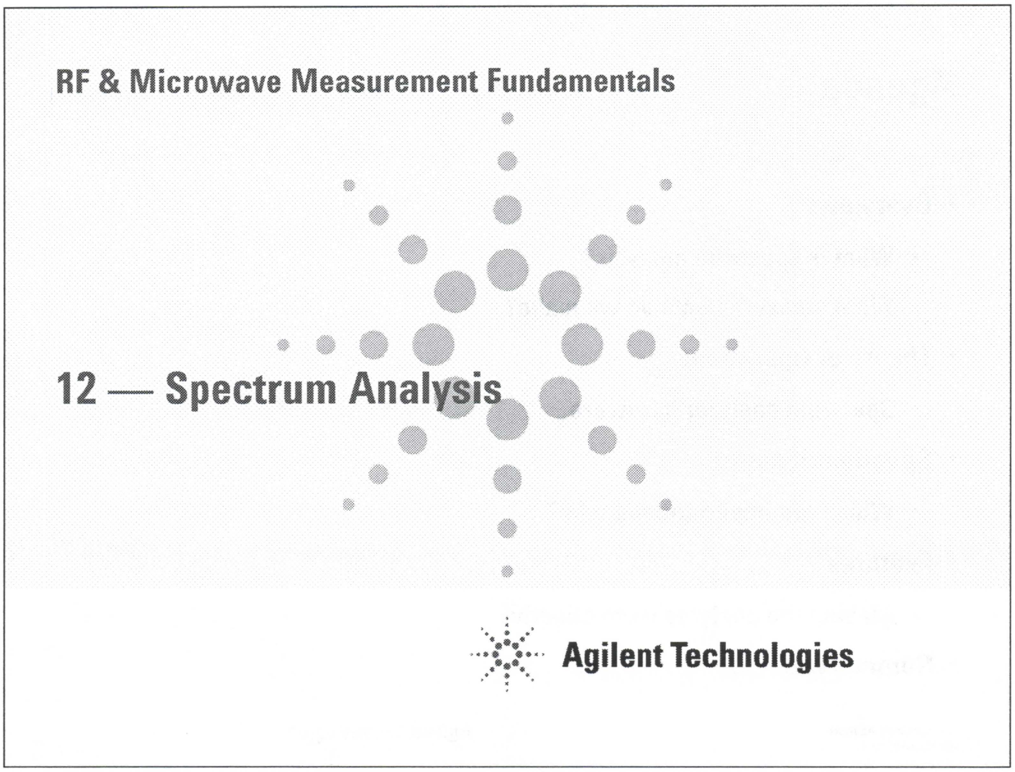


11 – Network Analysis
H7215A#101 v2.0

Agilent Technologies

Page 62

Here we see two traces, one showing the composite response and the second showing an improvement in filter rejection by removing the leakage and re-reflected responses.



This paper is intended to be a beginning tutorial on spectrum analysis. It is written for those who are unfamiliar with spectrum analyzers, and would like a basic understanding of how they work, what you need to know to use them to their fullest potential, and how to make them more effective for particular applications.

We will begin with an overview of spectrum analysis. In this section, we will define spectrum analysis as well as present a brief introduction to the types of tests that are made with a spectrum analyzer. From there, we will learn about spectrum analyzers in terms of the hardware inside, what the importance of each component is, and how it all works together.

In order to make measurements on a spectrum analyzer and to interpret the results correctly, it is important to understand the characteristics of the analyzer. Spectrum analyzer specifications will help you determine if a particular instrument will make the measurements you need to make, and how accurate the results will be. Spectrum analyzers also have many additional features that help make them more effective for particular applications. We will discuss briefly, some of the more important and widely used features in this section.

Agenda

- **Overview:**
  - **What is spectrum analysis?**
  - **What measurements do we make?**
- **Theory of Operation:**
  - **Spectrum analyzer hardware**
- **Specifications:**
  - **Which are important and why?**
- **Features**
  - **Making the analyzer more effective**
- **Summary**

12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 2

Objective

To get basic understanding of how spec an's work, how to use them to their fullest potential, and how to make more effective for particular applications.

Agenda

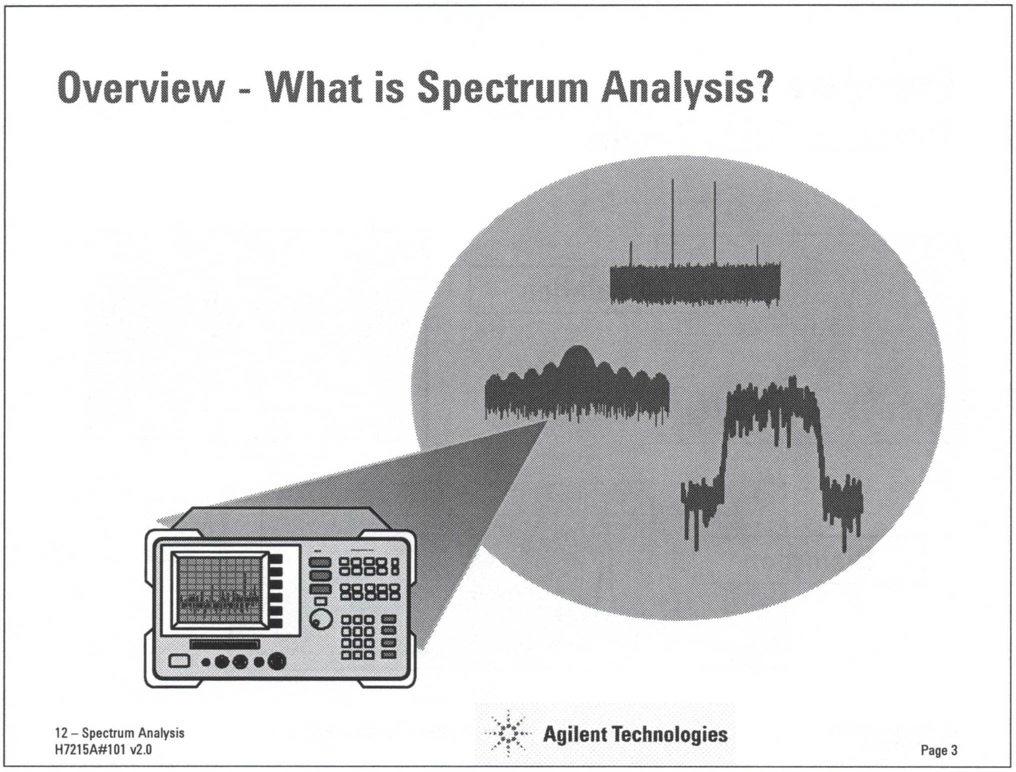
•Overview - to define spectrum analysis and introduce types of tests made with a spectrum analyzer

•Theory of Operation - to learn about the hardware inside an analyzer, and how the components all work together.

•Specifications - understanding the specifications of the analyzer will help determine if a particular instrument will make the required measurements and the accuracy of the results.

•Features - hardware and firmware features make the analyzer more effective for particular applications. We will discuss some of the more important and widely used features.

•Summary

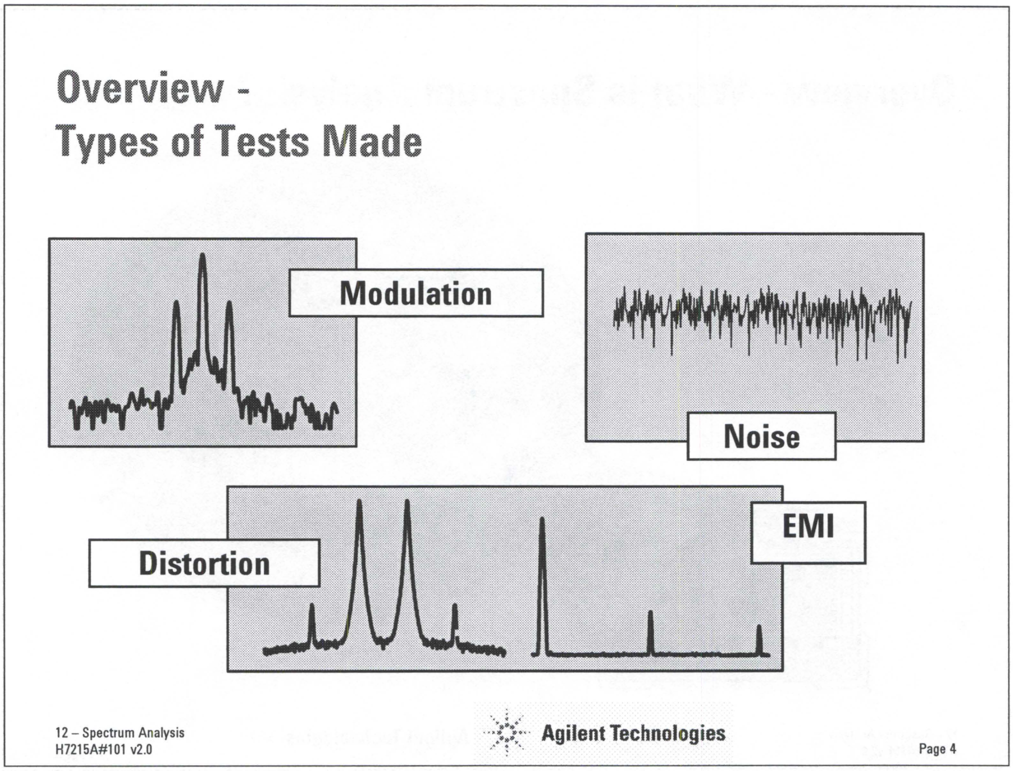


What is spectrum analysis?

Very basically, it is the analysis of the signal by finding its energy or power as a function of frequency. The spectrum analyzer is the tool of spectrum analysis, it allows viewing and measurement of the way energy is distributed in frequency. This is called the frequency domain.

By measuring signals in terms of their frequency domain components, insight into operation, performance specifications and troubleshooting are facilitated.

Understanding the important aspects of a spectrum analyzer will enable an operator make accurate measurements and to have confidence in the interpretation of the results.



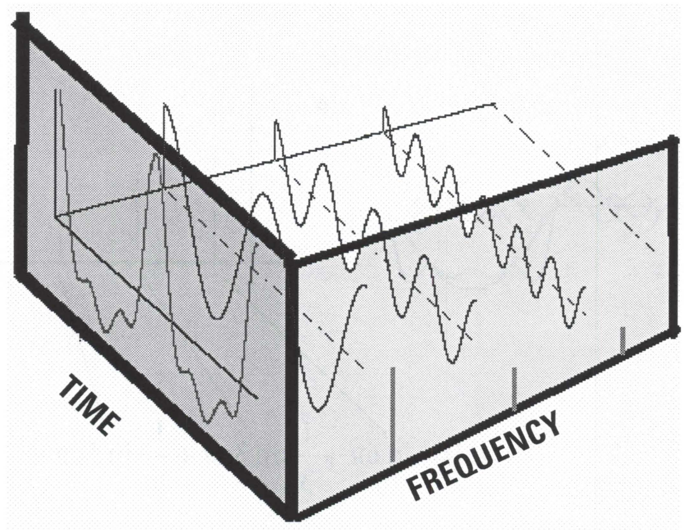
Some types of test

There are several different tests or measurements that can be made with a spectrum analyzer. The most usual are to do with modulation, distortion, or noise. When measuring a modulated signal; modulation degree, sideband amplitude, modulation quality, and occupied bandwidth are the usual parameters. The measurement of distortion of a system or device are necessary to verify performance specifications include: intermodulation, harmonics, and spurious emissions.

Not only is it important to understand the signal being transmitted, amplified, or filtered, but it is also very important to measure noise in the system ordevice in order to characterize its adverse effect on overall performance.

To comply with various national and international regulations, and for good engineering practice, the control of unwanted emissions is most easily understood with a spectrum analyzer. These are EMI, electromagnetic interference, measurements. Understanding the tests you need to make is critical for choosing the right measurement tool and getting the most out of it.

Frequency & Time Domain



12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 5

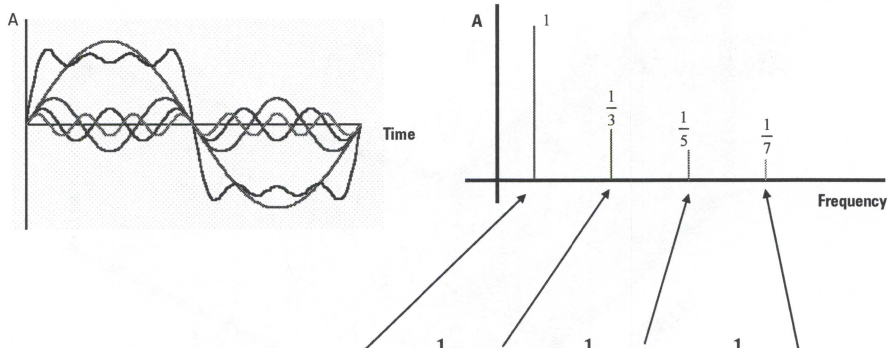
The well understood tool, used analyze an electrical signal, is an oscilloscope. It shows a graph of a signal's voltage as a function time. This is the Time Domain, most people understand the time domain and feel comfortable with time.

The spectrum analyzer is to the frequency domain as the oscilloscope is to the time domain, the frequency domain takes more thought to appreciate. Some things we do are both in the time and frequency domains, for example in listening to music there is a element that inhabits the time domain, one note follows another in a time progression but we hear a changing soundstage made of different tones; that is the frequency domain.

The figure shows a signal in both the time and the frequency domains. In the time domain, all frequency components of the signal are summed together and displayed. In the frequency domain, complex signals (that is, signals composed of more than one frequency) are separated into their frequency components, and the level at each frequency is displayed.

A signal that on an oscilloscope may appear to be a sinewave may not be perfect. On a spectrum analyzer, this lack of perfection shows as harmonics.

Wave Generation from Components



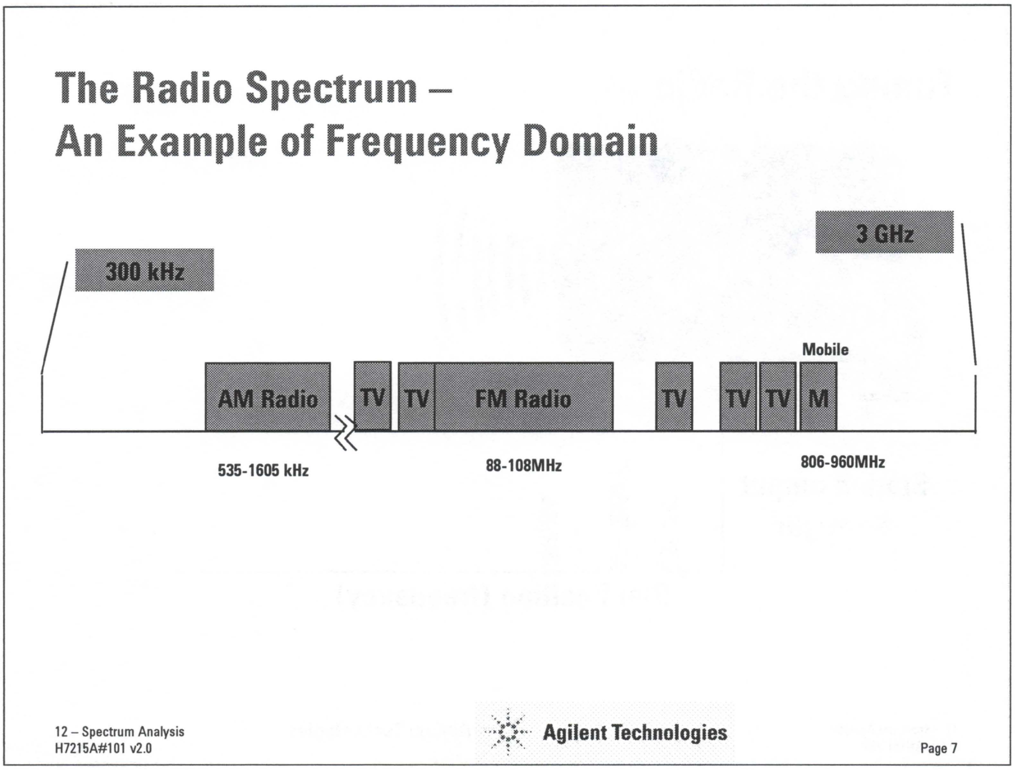
$$v(t) = \sin \omega t + \frac{1}{3}\sin 3\omega t + \frac{1}{5}\sin 5\omega t + \frac{1}{7}\sin 7\omega t + \ldots$$

12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 6

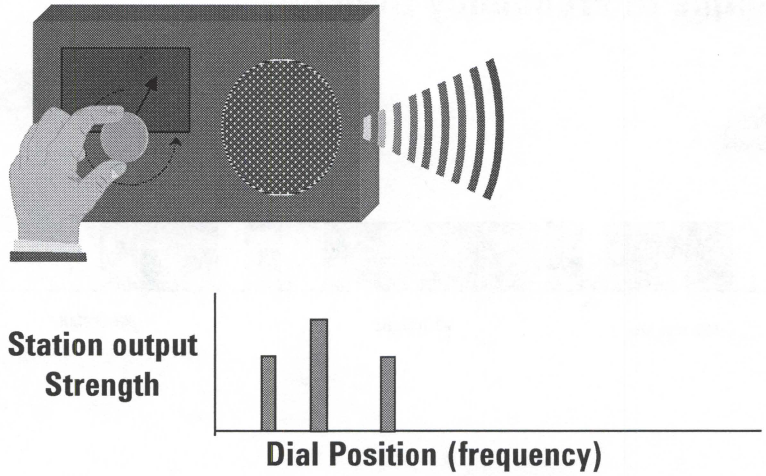
The frequency domain is usually displayed as a one-sided (positive frequencies only) power or amplitude spectrum. In this example the initial phase relationship between the components is zero degrees, this information is implicit in the time domain picture, but is lost in the amplitude spectrum. When a Fourier transform is done an amplitude and phase spectrum is generated, however in the swept spectrum analyzer the phase spectrum is not available.



Some systems are inherently frequency domain oriented. For example, many telecommunications systems use what is called Frequency Division Multiple Access (FDMA) or Frequency Division Multiplexing (FDM). In these systems, different users are assigned different frequencies for transmitting and receiving, such as with a cellular phone. Radio stations also use FDM, with each station in a given geographical area occupying a particular frequency band. These types of systems must be analyzed in the frequency domain in order to make sure that no one is interfering with radio stations on neighboring frequencies.

From this view of the spectrum, measurements of frequency, power, harmonic content, modulation, spurs, and noise can easily be made. Given the capability to measure these quantities, we can determine total harmonic distortion, occupied bandwidth, signal stability, output power, intermodulation distortion, power bandwidth, carrier-to-noise ratio, and a host of other measurements, using just a spectrum analyzer.

Tuning the Radio



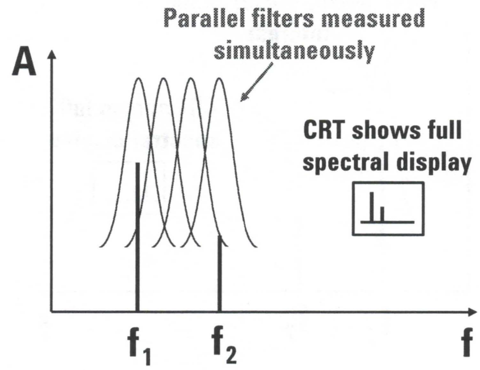
12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 8

Overview –
Different Types of Analyzers: Fourier

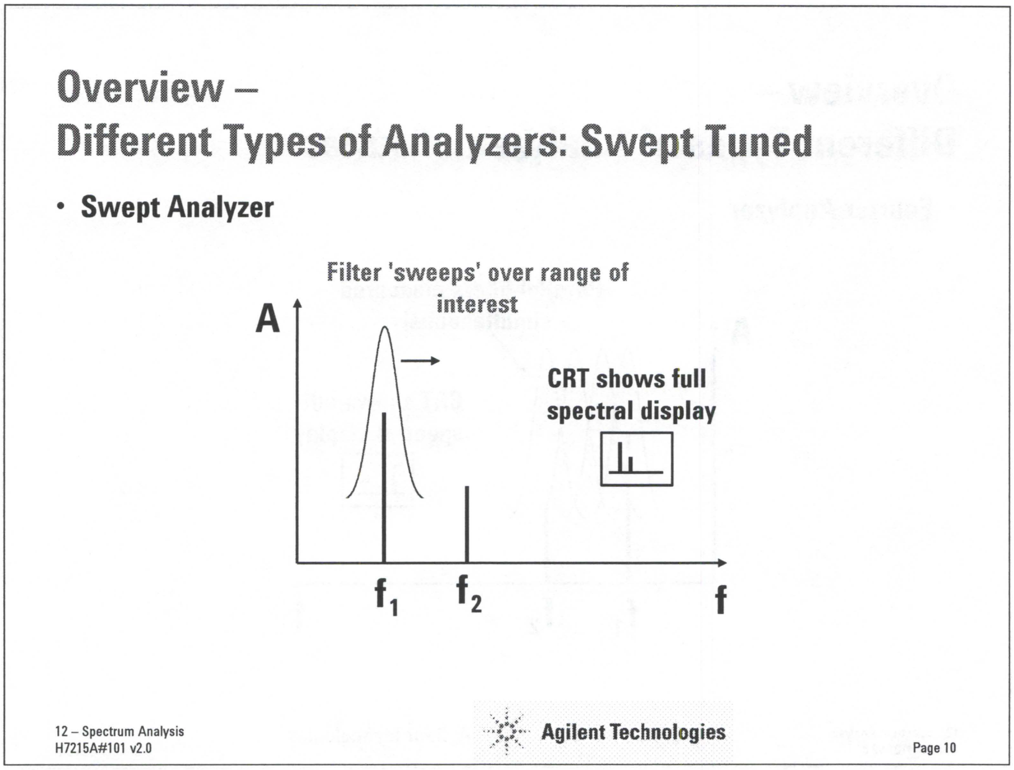
- **Fourier Analyzer**



The Fourier analyzer uses a time record of a signal, digitizes it using digital sampling, and then performs the mathematics required to convert it to the frequency domain, and display the resulting spectrum. It is as if the analyzer is looking at the entire frequency range at the same time using parallel filters.

With its real-time signal analysis capability, the Fourier analyzer is able to capture periodic as well as random and transient events. It also can provide significant speed improvement over the more traditional swept analyzer and can measure phase as well as magnitude. However it does have its limitations, particularly in the areas of frequency range, sensitivity, and dynamic range. We shall discuss what these terms are and why they are important in a later section.

As analog-to-digital converters (ADC) and digital signal processing (DSP) technologies advance, Fourier analyzers can operate at frequencies that are high enough to make them an important addition to the RF engineers toolbox. These analyzers can offer significant performance improvements over conventional spectrum analyzers, and will no doubt assume an increasingly important place as the analyzer of choice.



The most common type of frequency-domain analyzer is the swept-tuned receiver. Very basically, these analyzers "sweep" across the frequency range of interest, displaying all the frequency components present.

It works much like your AM radio, except that the dial controls the tuning on the radio, and the output is a speaker rather than a display. The present advantages of swept-tuned over Fourier analyzers are:

- •Wider frequency range
- •Larger dynamic range
- •Lower noise floor.

In most implementations of swept analyzers phase information is lost. In the next part of this lesson, the term spectrum analyzer will refer only to the swept-tuned analyzer.

Viewing the Frequency Domain with a Swept Analyzer

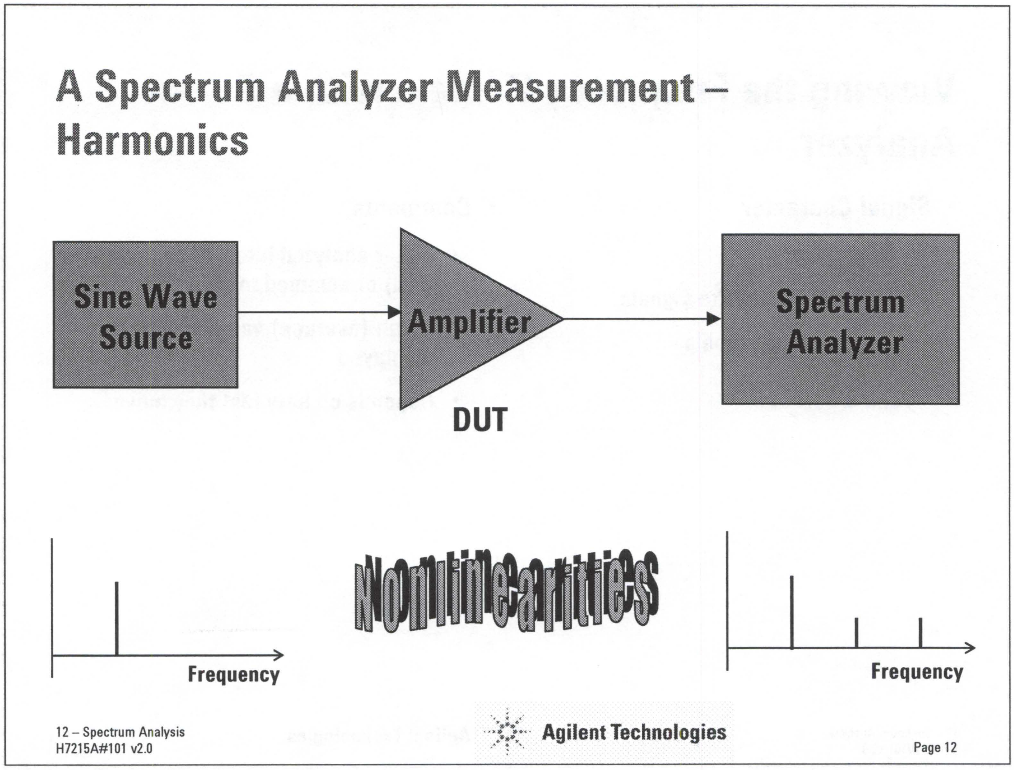
- **Signal Character**
  - **Repetitive**
  - **Noise & Noise Like signals**
  - **Time varying signals**
- **Comments**
  - **Either analyzed into CW components (NB) or summed to give (BB) display**
  - **Mean (average) value may be displayed**
  - **Depends on how fast they move**

12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 11

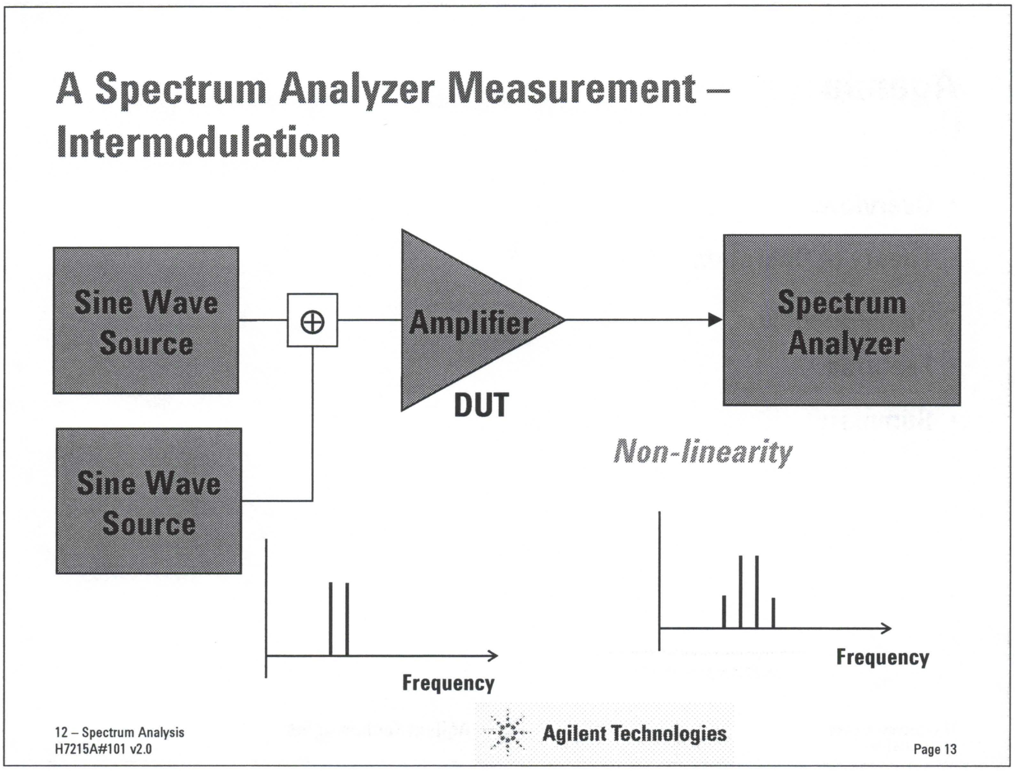
The display of the measurement of a received signal into the frequency domain depends on the nature of the signal. Since a finite time is taken (the sweeptime) to explore the received spectrum, the components of that spectrum must be substantially stationary from sweep to sweep. A signal that changes from sweep to sweep like noise, may be evaluated from its mean value. For other signal types that change from sweep to sweep this statistic may not be useful. Drifting signals may be watched provided they don't drift beyond the span boundaries from one sweep to the next.



Measurements of Nonlinearities

Unlike a Network Analyzer a Spectrum Analyzer can and does tune independantly from a source or signal stimulus. The spectral content of the DUT output may be viewed anywhere within the frequency range of the instrument. A major class of measurements done by a Spectrum Analyzer are the measurement of non linearities which are characterized by an output that is harmonicly related to the input stimulus.

It is these measurements that will be used as examples in future sections.



Intermodulation

This is a two tone test, two equal tones, f_0 and f_1 at a frequency separation Δf , are applied to the DUT. At the DUT output, extra signals at $f_1 + \Delta f$ and $f_0 - \Delta f$, and many others, are generated if the DUT has any third order coefficients in its transfer characteristic.

Harmonic and Intermodulation Measurements

These need sufficient *frequency range* to make measurements at both fundamental and harmonic frequencies and sufficient *resolution* to discriminate between the carrier(s) and the third order components for the two tone measurement.

Agenda

- Overview
- Theory of Operation
- Specifications
- Features
- Summary

12 – Spectrum Analysis
H7215A#101 v2.0

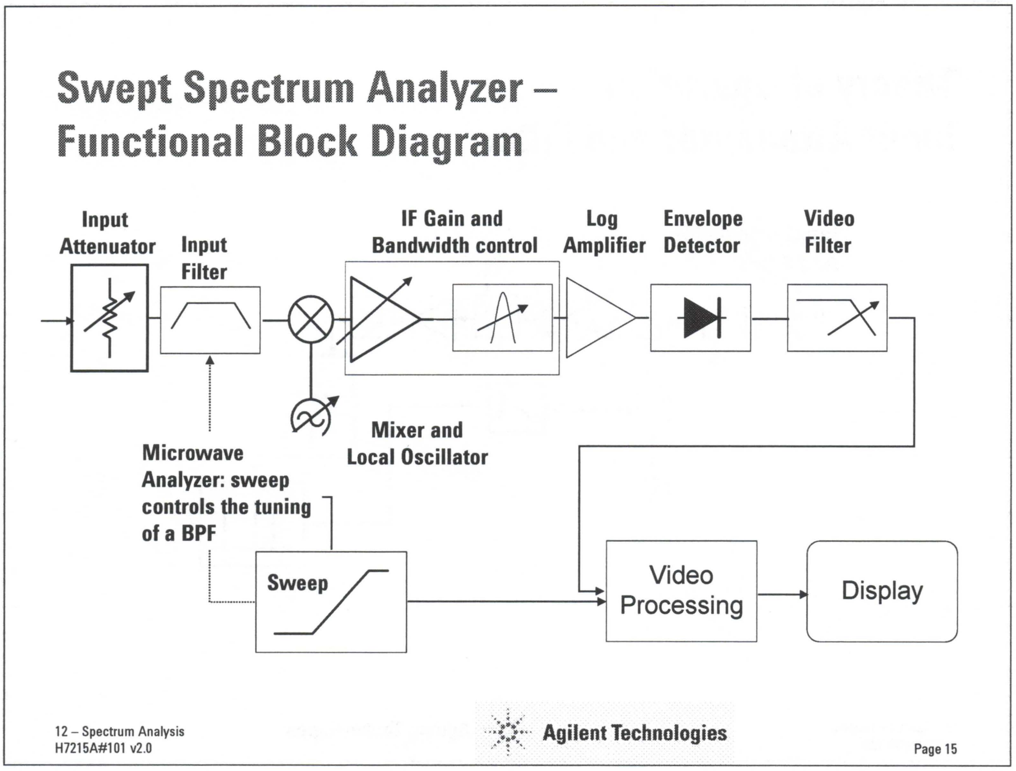
Agilent Technologies

Page 14

First; the theory of operation of the swept tuned spectrum analyzer.

The function of a spectrum analyzer may be explained by imagining a band pass filter which is tuned across the frequency band of interest, as signals pass through the filter, they are detected and displayed. This is a good elementary picture and has been used at microwave frequencies where a *qualitative* view of the frequency domain is all that is needed. Filters that tune over a few octaves at microwave frequencies are available, they are called YIG (Yttrium Iron Garnet) filters. 500MHz to 2 GHz; 2 to 16GHz etc. This is not a practical solution for an instrument to give *quantitative* results, nor are there filter technologies that would make an economic full range (e.g. 10kHz to 3GHz) analyzer possible.

A practical solution is the use of the heterodyne principle as used in radio receivers, here the *spectrum is swept across a fixed filter*.



Block Diagram

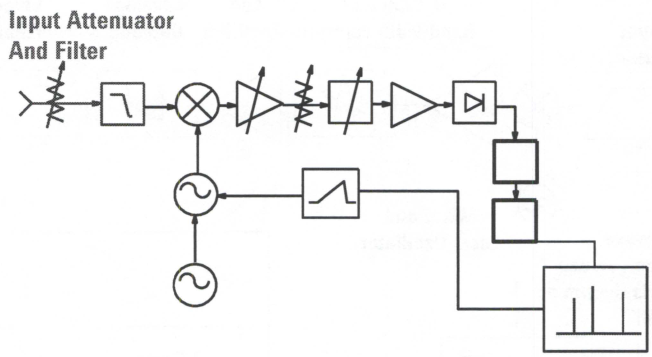
This is the basic block diagram for the swept-tuned spectrum analyzer. The analyzer will have a much more complicated block diagram than this but from the point of view of the functioning of the analyzer this diagram works well.

Before considering how it all works together to create a display of frequency versus amplitude on the screen, the major functional components will be briefly discussed.

Why is it important to understand the block diagram?

To some extent, even for simple instruments, knowing how the instrument works can add confidence for the user and authority to the results. A spectrum analyzer is not a simple instrument and may be used for many applications and types of signals. Except for the most simple measurements with a spectrum analyzer some degree of expert knowledge is necessary for good measurement practice and interpretation of results. A knowledge of the block diagram will go a long way towards that expert knowledge.

Theory of Operation – Input Attenuator and Filter



12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

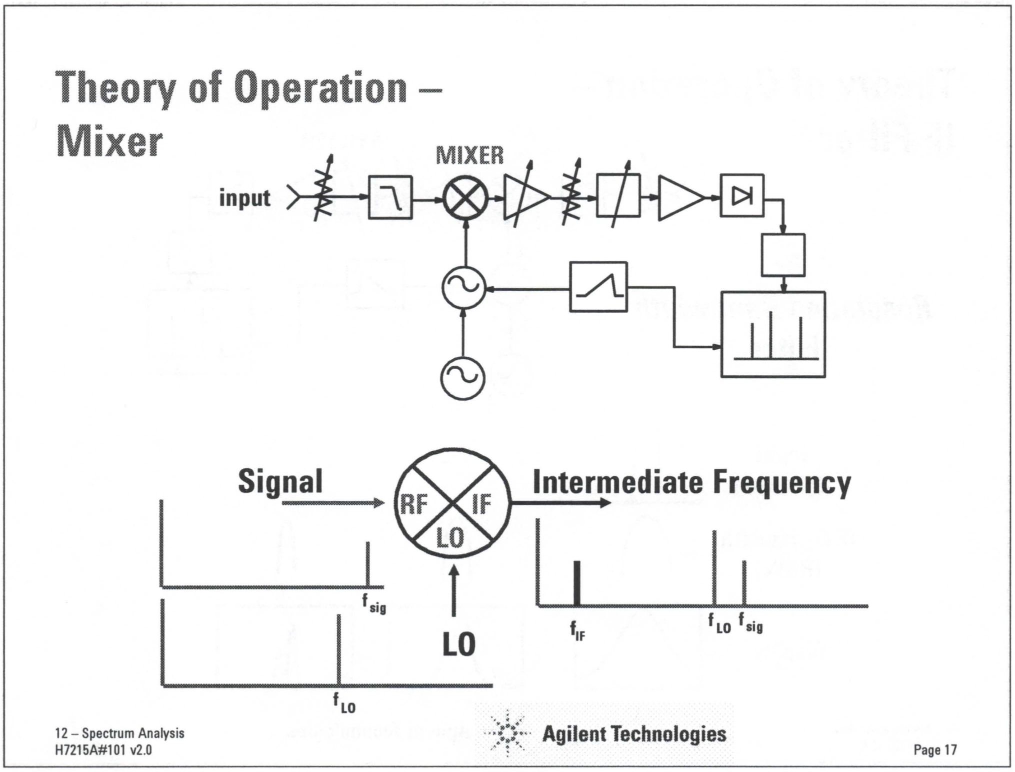
Page 16

Input Attenuator

The RF input attenuator is used to adjust the level of the RF signal into the mixer. It is important to be able to do this in order to protect the mixer from large signals and to control the spurious free dynamic range of the spectrum analyzer. The RF attenuator may be set to zero dB, however this condition must be used with care, analyzer designers have used mechanical interlocks or special button sequences to avoid an operator accidentally selecting 0dB attenuation. Most analyzers have attenuators with the range 0dB to 70dB with ten or five dB step selection.

Input Filter

This limits the amount of RF energy to the band of the analyzer. In an RF analyzer this band is the whole RF bandwidth of the instrument. In a microwave spectrum analyzer (for example > 3GHz) this filter is a tunable bandpass filter which tunes with the sweep, called a microwave preselector.



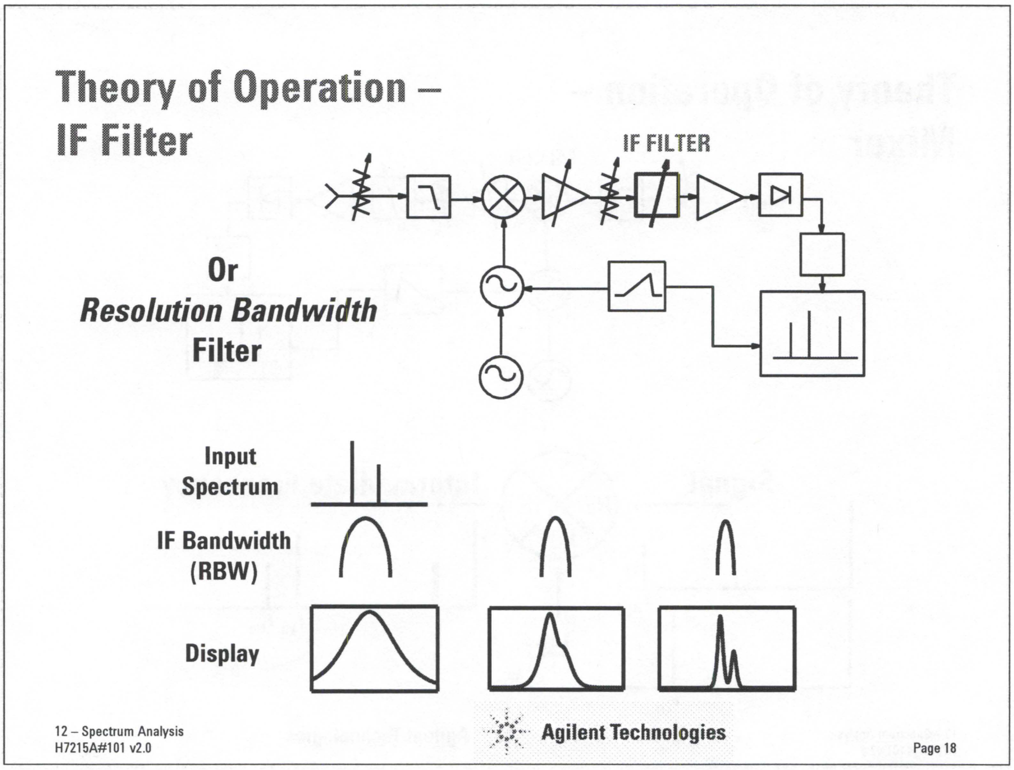
The Mixer

There are many mixing stages in a spectrum analyzer, we can explain the function of the analyzer by treating it as a simple one stage receiver. Mixers are three-port devices that convert a signal from one frequency to another, the mixer may be called, a frequency translation device. The RF input signal is applied to the input port, and the Local Oscillator signal to the L port. The LO is a high level signal (> +10dBm) compared to the RF (< –10dBm). A mixer is a non-linear device, this means that besides the wanted frequency translated signal, there are a number (theoretically an infinite number) of signals we don't want. The wanted signals from the mixer are (fR – fL) or (fL – fR), positive quantities. It is this difference frequency that is of interest in the spectrum analyzer, We call this signal the IF signal, or Intermediate Frequency signal.

Microwave Spectrum Analyzers

The IF is a relatively low frequency, so the LO must generate frequencies over the same nominal frequency range as the analyzer's input frequency range. In microwave spectrum analyzers, therefore, to avoid the expense of providing a fundamental LO which would be a wide range microwave LO, multipliers may be used or more usually harmonic relationships at the mixer output are used.

- (fR –n.fL) or (n.fL – fR)



The Intermediate Frequency (IF) or Resolution Bandwidth (RBW) Filter

The IF filter is also known as the Resolution bandwidth filter. When you change the RBW on the analyzer, you change the bandpass width of this filter.

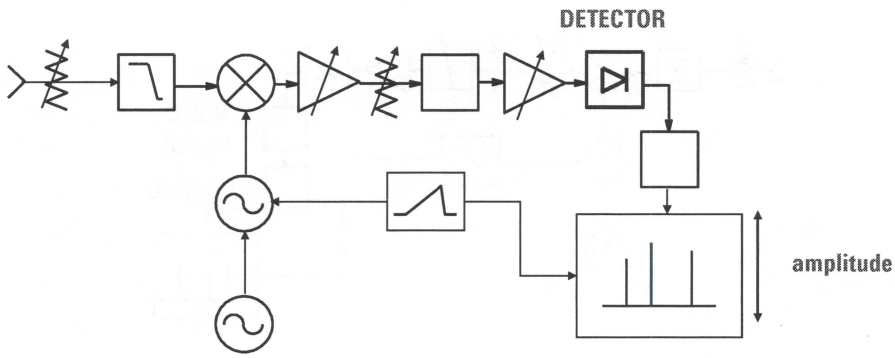
The IF filter is a bandpass filter that is used as the "window" for detecting signals.

Spectrum analyzers typically give several choices of RBW settings. By having a broad range of variable RBW settings, the analyzer can be optimized for a given measurement condition.

The example in this illustration shows that as the filter is narrowed, selectivity is improved and two closely spaced input signals may be resolved. This will, however, slow down the sweep speed.

The resolution bandwidth filter will be discussed in other parts of this class.

Theory of Operation – Detector (Envelope Detector)



12 – Spectrum Analysis
H7215A#101 v2.0
Agilent Technologies
Page 19

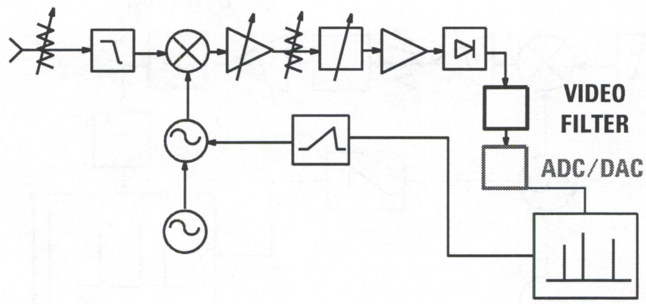
The IF Signal Detector

The detector converts the IF signal to a baseband or video signal so it can be viewed on the instrument's display. It is the envelope of the IF signal which is proportional to the input RF spectrum amplitude, so the detector is called an envelope detector.

It is the output of this envelope detector which is needed to deflect the trace in the y-axis, or amplitude axis, of the display. (In older spectrum analyzers the waveform from the envelope detector would be amplified and applied to the y-axis of a CRT display.)

In spectrum analyzers using analog CRT displays, great efforts were made to maintain the trace on the screen especially during long sweeptimes, so called storage CRT's were employed, which took much practice to adjust properly.

Display Detection Process



12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

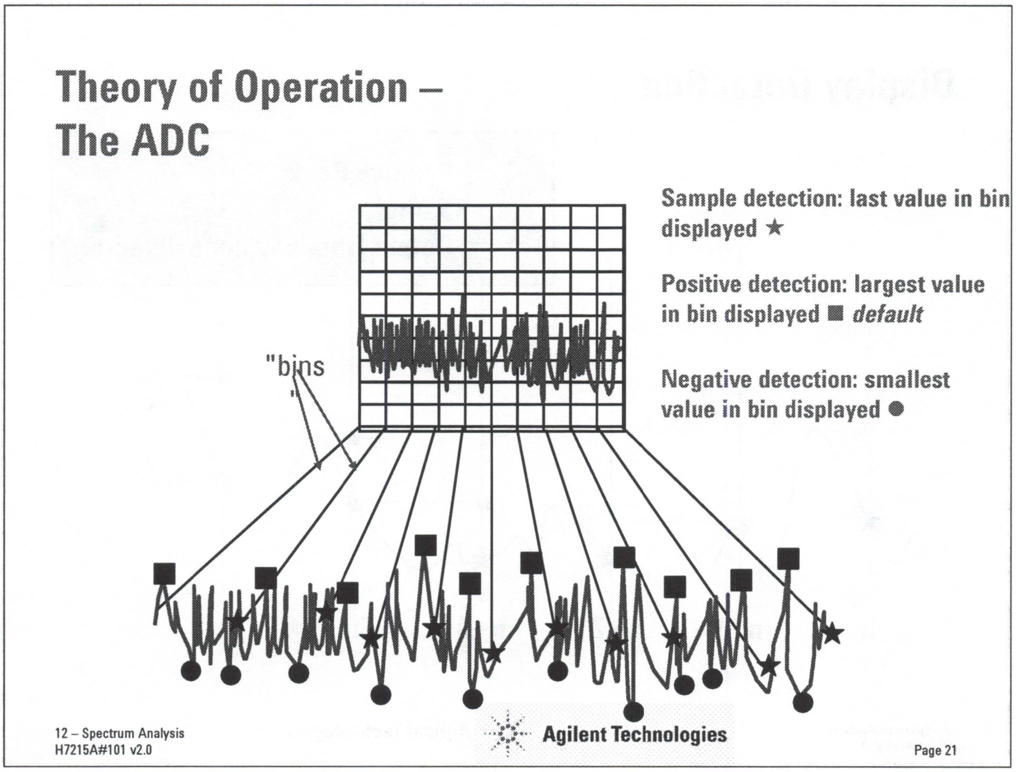
Page 20

Before we discuss the video filter, a brief look at the video processing is in order. This is because the display of a video filtered signal is affected by the type of digitization selected.

This block called ADC/DAC may be adjusted from the front panel by the selection of such functions as +ve peak, or sample detection, or by max hold and min hold. The term **display detection** used here must not be confused with the envelope detection process.

It is the analog output of the envelope detector which is the waveform to be subject to digitization.

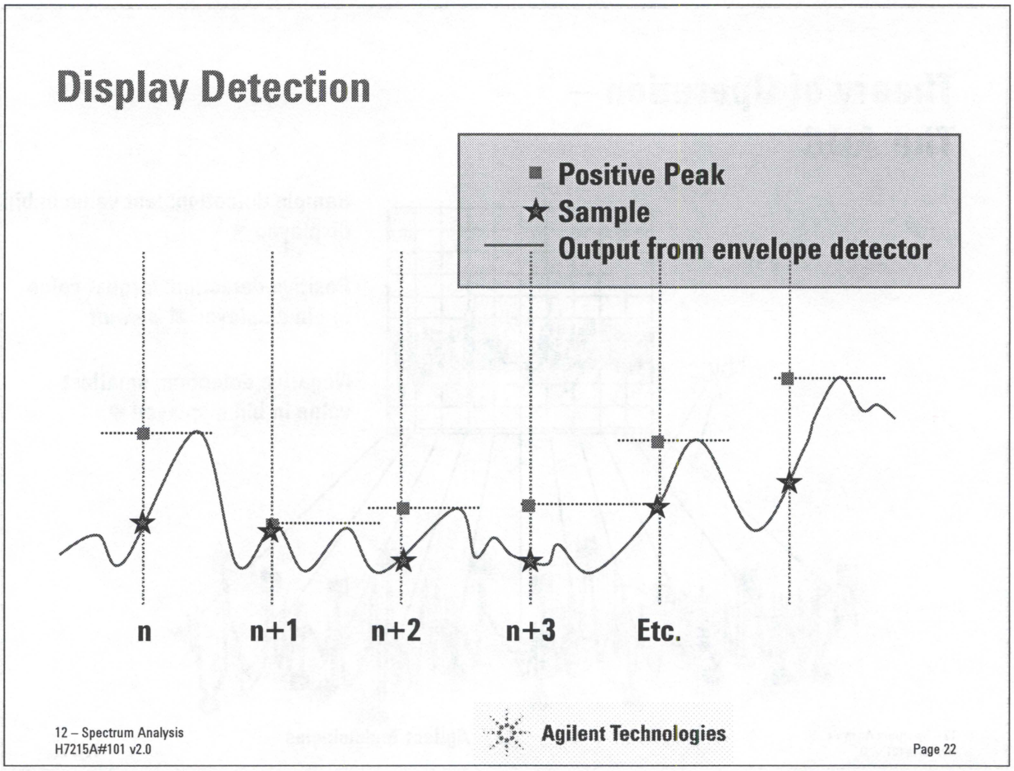
The digitized values are stored, then replayed to the CRT at the refresh rate via the DAC which effectively "connects the dots" to create the frequency domain graph for viewing. The way this digitization is done will dramatically effect how the signal is displayed.



The x-axis of the display can be thought of as being made up of "bins" (or trace elements) from which the analog detector output is digitally sampled. A natural question is: what point in the bin do we use for our data point?

In positive peak detection mode, the peak value of the signal over the duration of one trace element is used, whereas in negative peak detection mode, it's the minimum value. Positive detection mode is typically used when analyzing sinusoids,(CW and stationary signals), but is not good when noise, or noise-like measurements are needed, because the value of the noise level is biased positively by the positive peak process.

In sample detection, since the sample is taken at evenly spaced time intervals, there is no bias in the sampled data, so the statistics of noise and noise like signals are preserved. This detection mode is used for measuring noise or noise-like signals.


Display Detection
Positive Peak
Sample
Output from envelope detector
n
n+1
n+2
n+3
Etc.
12 – Spectrum Analysis
H7215A#101 v2.0
Agilent Technologies
Page 22

A/D Detection modes

- **Positive peak detection:**
  - **Largest value in bin is entered, so never misses a signal. Used for stationary signals like CW. This is the default mode of most modern swept analyzers.**

- **Sample detection:**
  - **Last value in bin is entered, this means that the sampling is evenly spaced in time. Used for measurements of noise and noise-like signals.**

12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 23

A/D Detection modes (II)

- **Negative peak detection:**
  - **Smallest value in bin is entered, used for the min-hold . -ve peak and +ve peak may be used together for special display modes.**

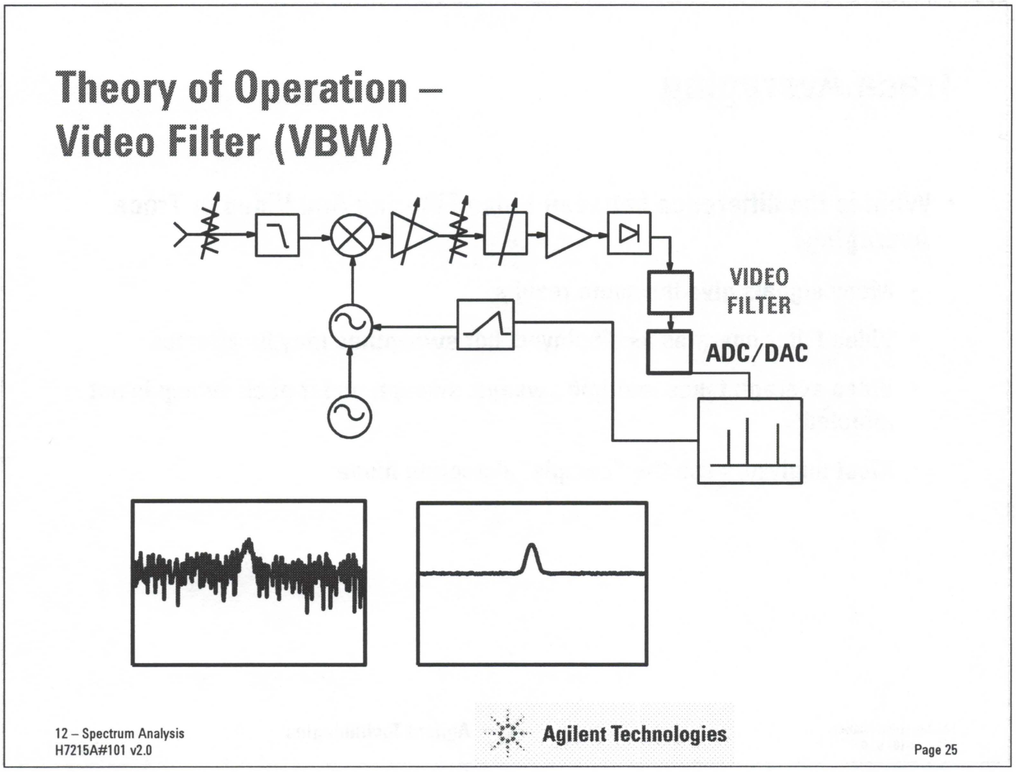
12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 24

In some Agilent spectrum analyzers the default mode uses a special algorithm which has been given the name "rosenfell"* mode. This mode is designed to display noise without bias and CW signals properly. For example, if the signal both rose and fell within a sampling bin, it assumes it is noise and will use the positive & negative detector alternately. If it continues to rise, it assumes a CW signal and uses positive peak detector.

* The name is derived from " if the data rose and fell, then . . ."



The video filter is a low-pass filter that is located after the detector and before the ADC.

It is used to average or smooth the trace that is seen on the display, as shown on the slide. As the video filter BW is decreased the trace will become smoother.

- A change in video BW does not change the mean noise level.

The small decrease in displayed noise level observed, when the VBW filter is reduced and the positive peak detector is in effect, is because the displayed noise is biased above the mean by the +ve peak detection process.

The smoothing of the displayed noise allows signals close to the noise level to be observed. Remember that the displayed trace represents signal + noise, even when the noise is smoothed.

Trace Averaging

- **What is the difference between Video Filtering And Video or Trace Averaging?**
  - **Many signals give the same results**
  - **Video Filter operates as displayed, but sweeptime may be affected**
  - **Video average takes multiple sweeps, sweeptime for each sweep is not affected**
  - **Most analyzers use the "sample" detection mode**

12 – Spectrum Analysis
H7215A#101 v2.0

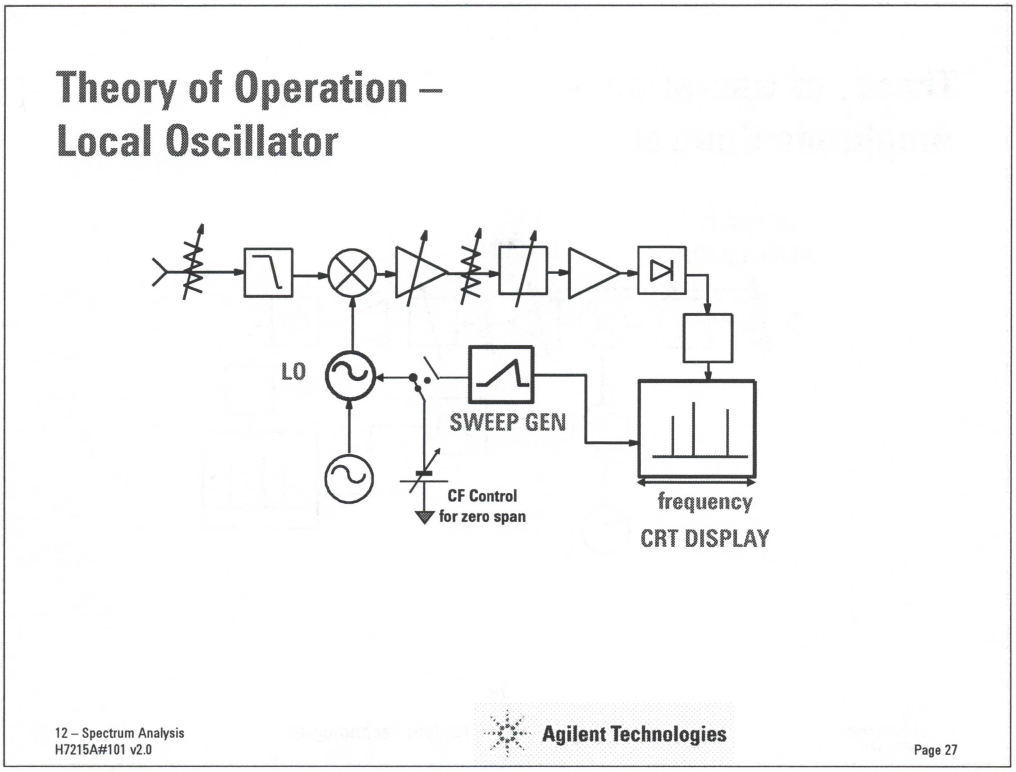
Agilent Technologies

Page 26

Video or trace averaging is a feature of most modern analyzers, in this case averaging is performed over two or more sweeps on a point-by-point basis. At each display point the new value is calculated from the new measured value and the previously averaged data.

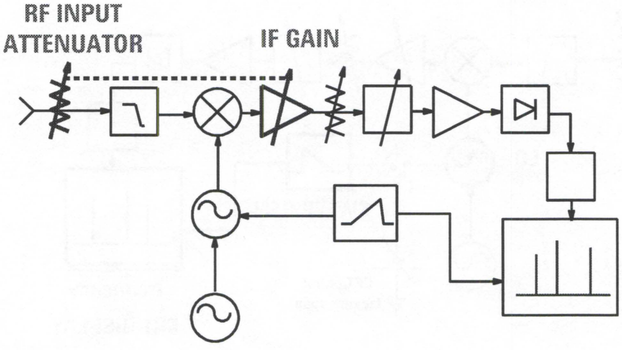
$$A_{avg} = \left[\frac{(n-1)}{n}\right].A_{prior\ avg} + \left(\frac{1}{n}\right).A_n$$

- Aavg = new average value
- Aprior avg = average from prior sweeps
- An = measured value on current sweep
- n = number of current sweep



The LO is a voltage controlled oscillator (VCO) which provides the LO signal for to the mixer. A sweep generator provides the ramp voltage to tune the LO in proportion to the ramp voltage. It also controls the horizontal position of the trace displayed on the spectrum analyzer display. This makes the display X axis the frequency axis. The ramp voltage may switched from the LO while still sweeping the display, in this mode the analyzer behaves like a fixed tuned reciever. This is called zero span or zero scan, the tuned frequency may be changed like changing stations on an AM radio.

Theory of Operation – Amplitude Control



12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 28

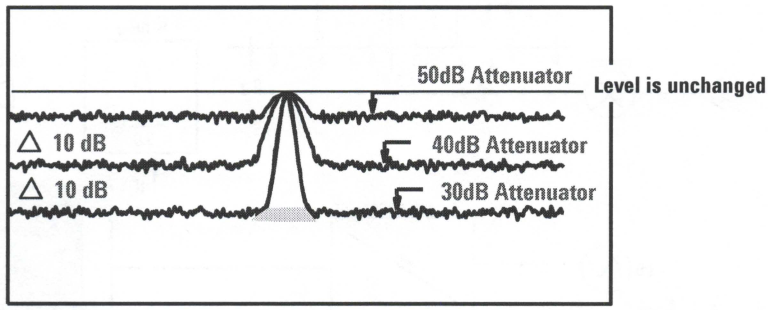
The IF gain is used to adjust the vertical position of signals on the display without affecting the signal level at the mixer. When we change this level, the reference level is changed accordingly.

The RF attenuator value and IF amplifier gain, are coupled. When the RF input attenuator is changed, the IF gain will automatically change so that signals will remain stationary on the screen.

Sweep time K $\frac{SPAN}{RBW^2}$

Theory of Operation – Attenuator/IF Amplifier Gain

- **Displayed Noise is a Function of Attenuator Setting**



Increased Attenuation = Decreased Signal/Noise ratio.

12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 29

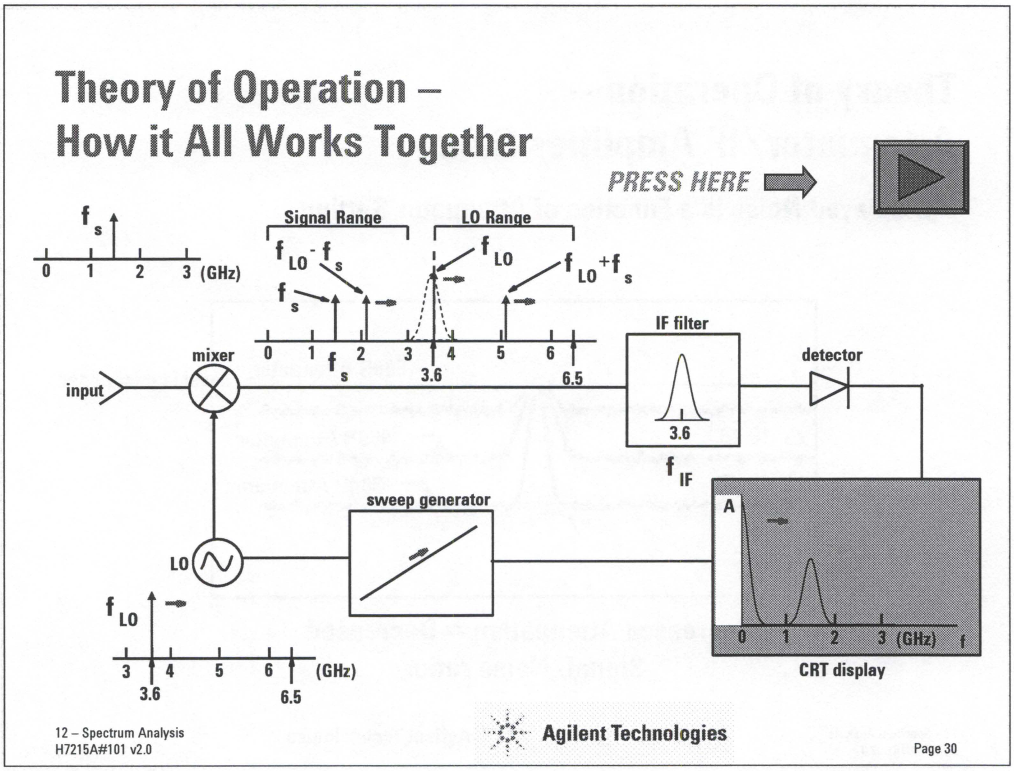
The Displayed Average Noise Level, DANL, is also a function of RBW

The internally generated noise in a spectrum analyzer is random and has no discrete spectral components. Also, its level is flat over a frequency range that is wide compared to the RBW ranges.

This means that the total noise reaching the detector (and displayed) is related to the RBW selected.

Since the noise is random, it is added on a power basis, so the relationship between displayed noise level and RBW is a ten log basis.

In other words, if the RBW is increased (or decreased) by a factor of ten, ten times more (or less) noise energy hits the detector and the displayed average noise level (DANL) increases (or decreases) by 10 dB.

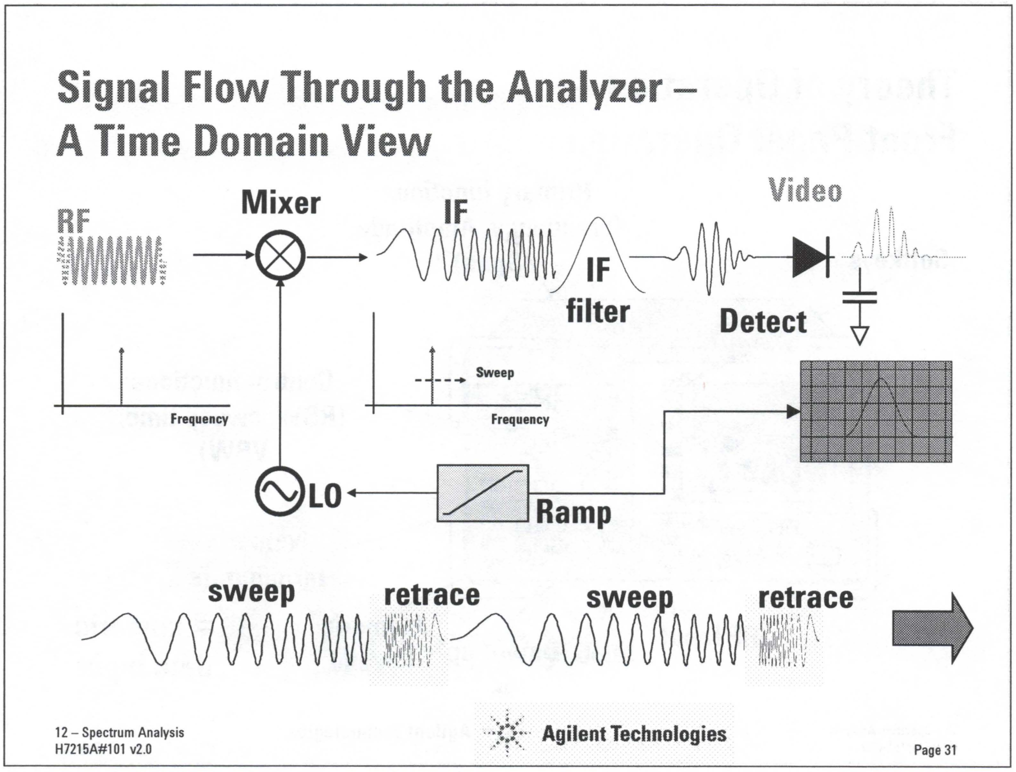


Note that while the RF input attenuator, IF gain, and video filter are important for understanding the final measurement, they are not critical when describing the signal flow.

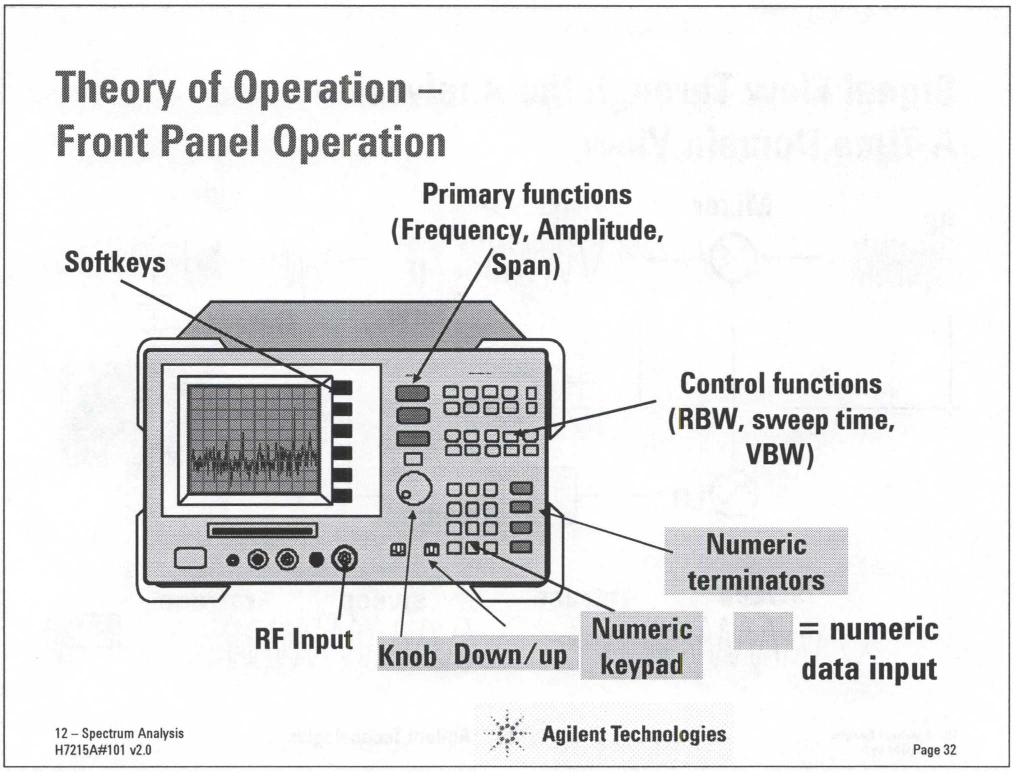
The signal to be analyzed is connected to the input of the analyzer. This signal is then combined with the LO in the mixer to convert it to an IF. The part of the IF signal that corresponds to the IF filter passband is detected, this detected output will be proportional to the amount of IF signal in the passband.

The output voltage of the detector drives the vertical axis (amplitude) of the CRT display.

The sweep generator output tunes the LO, while controlling the pen position on the display horizontal axis. The resulting display shows amplitude versus frequency of the spectral components of each incoming signal.



This is a time domain view of the signal flow through the analyzer. Some aspects such as the way the display shows the IF filter shape will be more clear with this view. Notice how the envelope detector will have a repetitive IF pulse when the signal is CW.



For those of you that have used spectrum analyzers before, you will already have an understanding of the front panel operation. Each hardkey reveals a softkey list of related functions activated by the buttons next to the display area. The three main functions used when setting up the analyzer are:

- Frequency (Center Frequency) - where in frequency spectrum is the analyzer tuned?
- Amplitude (Reference Level) - What is the full scale of the display?
- Span - how big is the range of frequency to view? These are usually the largest hardkeys on the front panel.

The other main control functions are typically smaller keys. Controlling RBW, sweep time, and VBW are a few of the more important functions.These functions are automatically controlled, when using the three main functions after a preset or switch on. Pressing a hardkey will display "softkeys"which provide detailed or related access to the hardkey function.

Most analyzers allow you to enter values by either punching in the value on the numeric keypad, dialing up or down to the desired value using the front panel knob, or by stepping up or down a given amount using arrow keys. Data entered by the numeric keypad must be terminated with a unit key.

Agenda

- **Overview**
- **Theory of Operation**
- Specifications
- **Features**
- **Summary**

12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 33

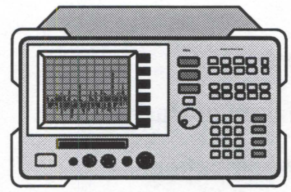
Understanding the capabilities and limitations of a spectrum analyzer is a very important part of understanding spectrum analysis. Today's spectrum analyzers offer a great variety of features and levels of performance. Reading a datasheet can be very confusing. How do you know which specifications are important for your application and why?

Spectrum analyzer specifications are the instruments manufacturer's way of communicating the level of performance you can expect from a particular instrument. Understanding and interpreting these specifications enables you to predict how the analyzer will perform in a specific measurement situation.

We will now describe a variety of specifications that are important to understand.

Specifications

- Safe Power Levels etc.
- Frequency Range
- Resolution
- Sensitivity
- Distortion
  - Dynamic Range
- Accuracy, Frequency & Amplitude



12 – Spectrum Analysis
H7215A#101 v2.0

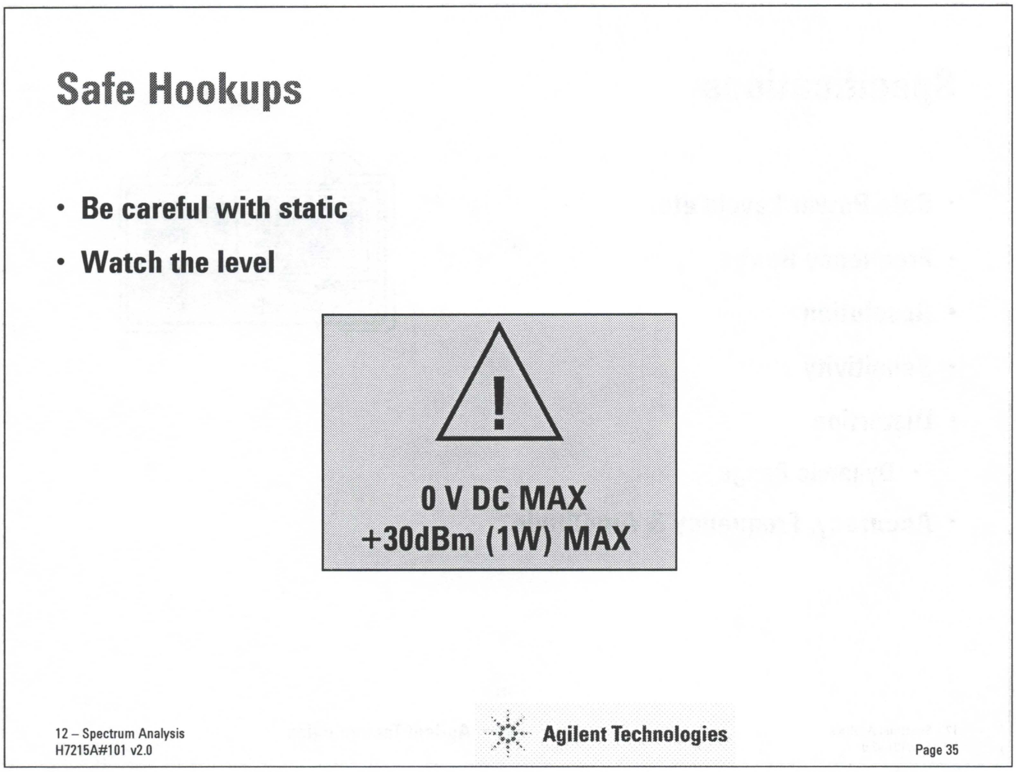
Agilent Technologies

Page 34

What needs to be known about a spectrum analyzer in order to make sure that it will make particular measurements, and make them adequately? The most basic performance parameters are:

- What is a safe power level?
- What's the frequency range?
- What's the amplitude range (maximum input and sensitivity)?
- How different in amplitude can two signals be simultaneously present and be measured accurately?
- How accurate are the measurements made with a spectrum analyzer?

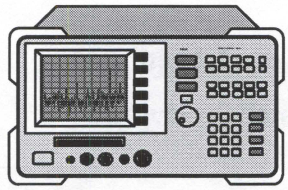
Although not in the same order, we will describe each of these areas in terms of what they mean, and why they are important.



Before connecting the signal to a spectrum analyzer (or any instrument) be sure that there is no charge on the cable and be aware of input limitations. These are usually printed close the terminals.

Static precautions are usually observed very strictly in production environments and should be taken seriously in less structured situations. Although the effect of static discharge may be obvious if it destroys the instrument input, often the effect is gradual, causing a progressive deterioration in performance.

Specifications

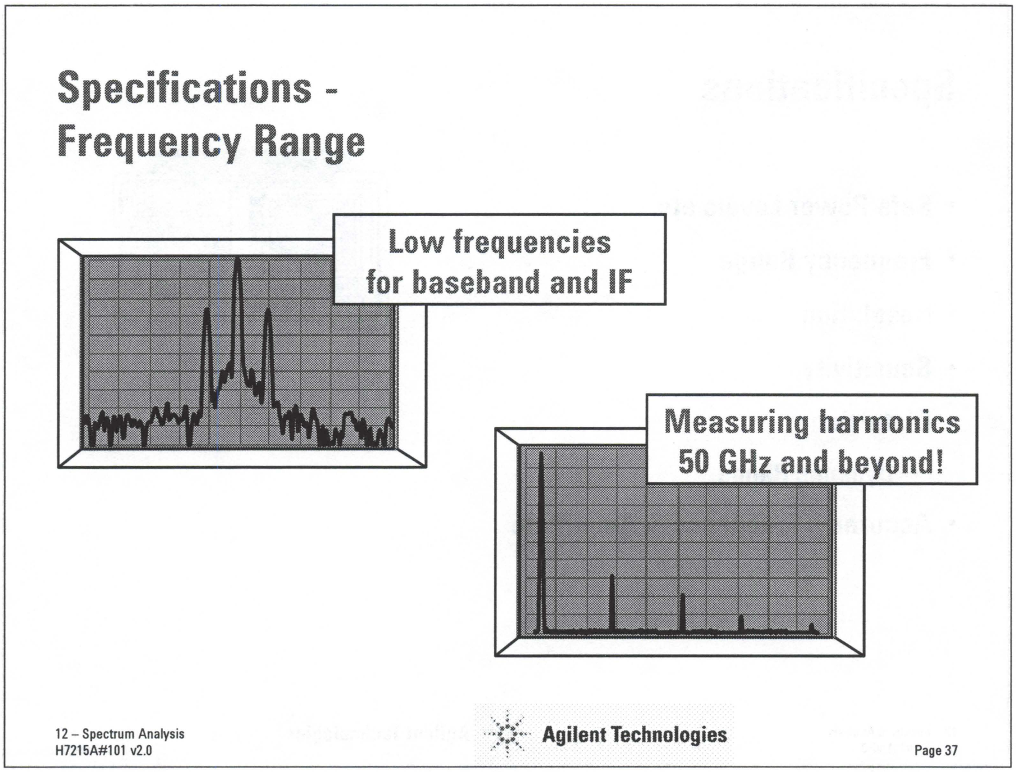


- Safe Power Levels etc.
- Frequency Range
- Resolution
- Sensitivity
- Distortion
  - Dynamic Range
- Accuracy, Frequency & Amplitude

12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 36



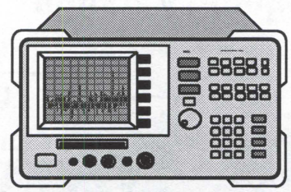
Frequency Range

The spectrum analyzer must measure not only the characteristics of the wanted signal but also undesired effects such as harmonics an spurs.

For example, some electromagnetic compatibility specifications (EMC) for communications systems, require measurements to the tenth harmonic. If the system operates at 900 MHz, then measurements to 9 GHz must be made. For such a system measurements at lower frequencies, such as baseband and IF, would also be useful.

Specifications

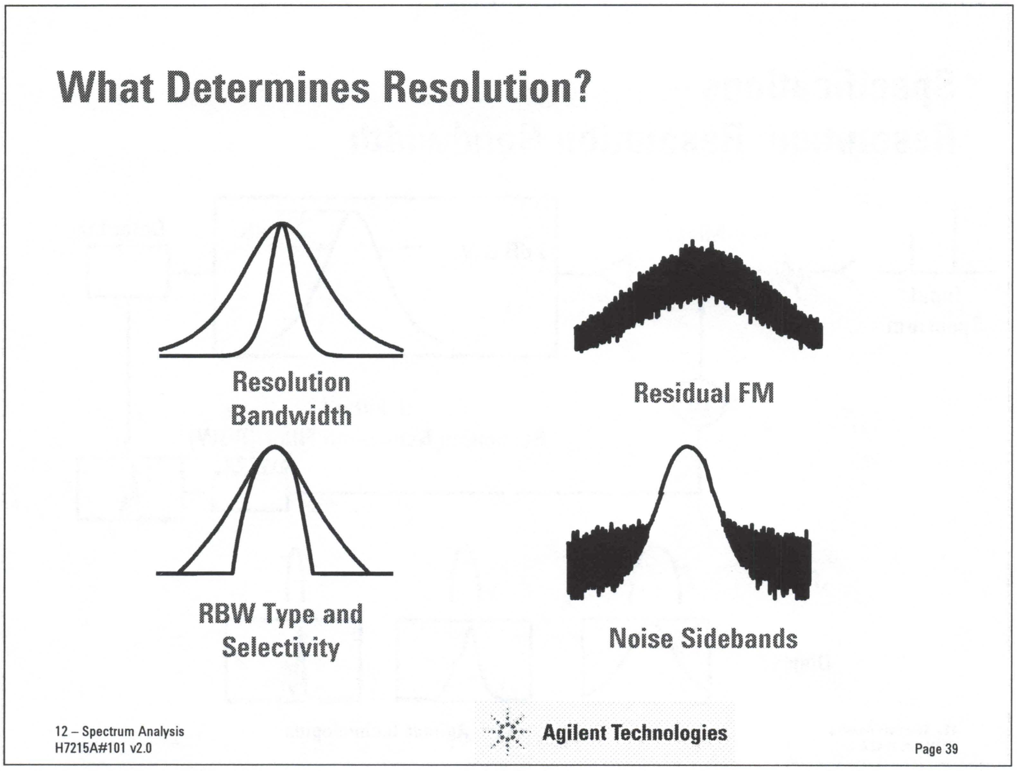
- Safe Power Levels etc.
- Frequency Range
- Resolution
- Sensitivity
- Distortion
  - Dynamic Range
- Accuracy, Frequency & Amplitude



12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

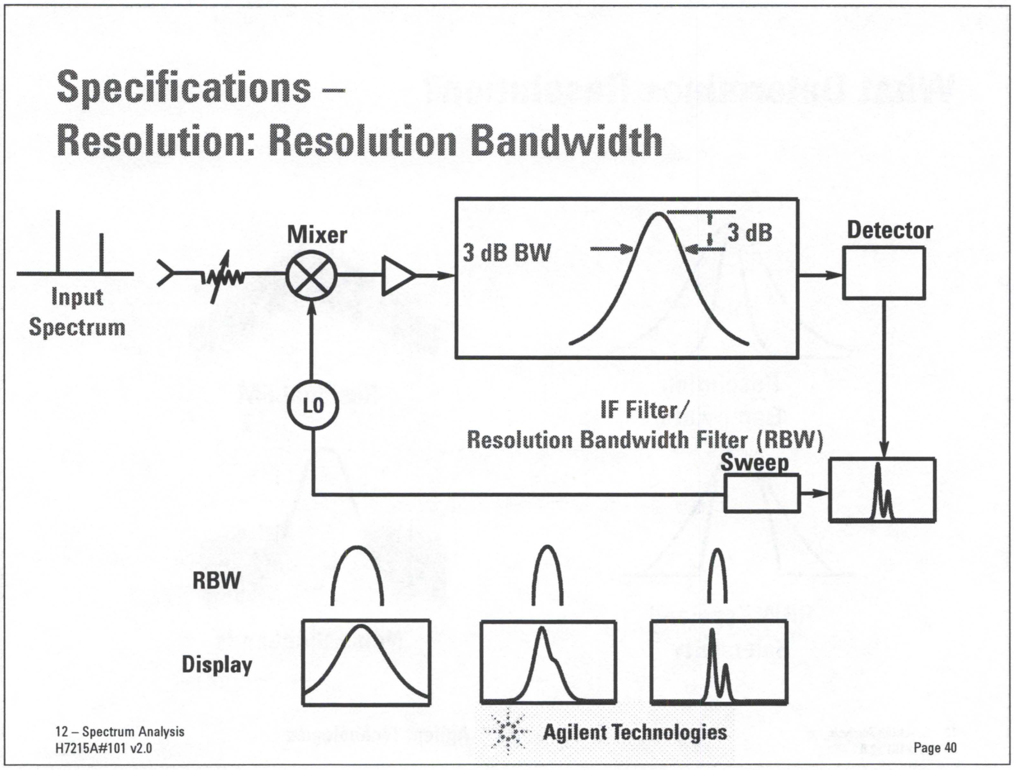
Page 38



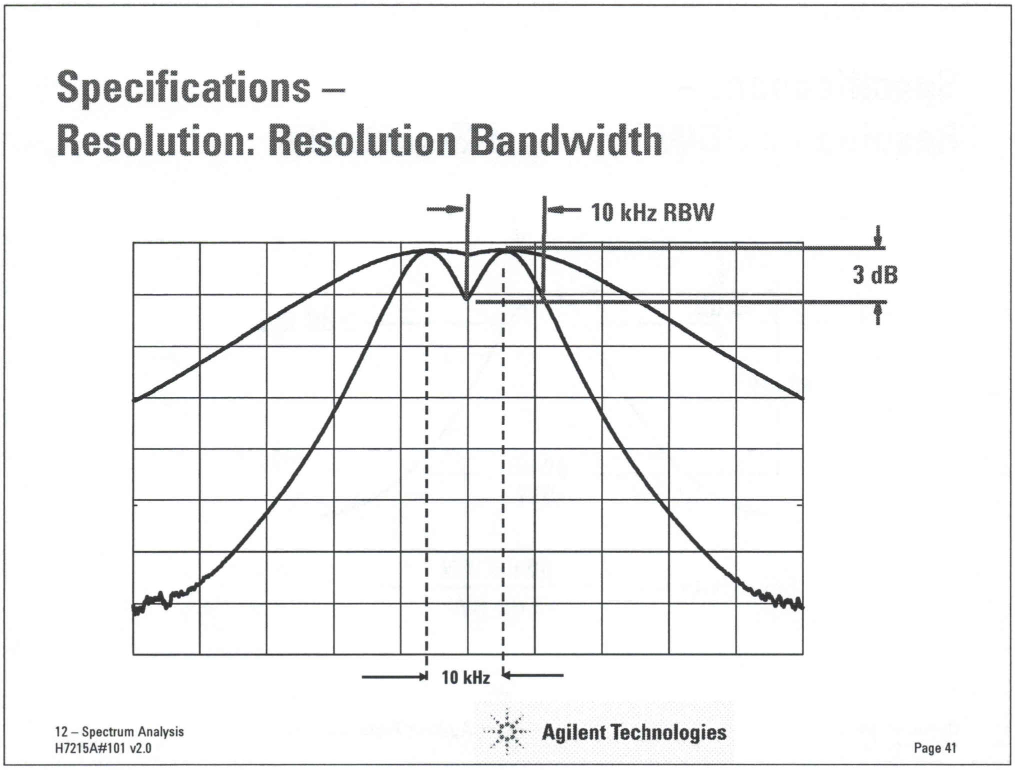
Resolution

This is an important specification when trying to measure signals that are close together. The IF filter is also called the resolution bandwidth(RBW) filter. This is because it is this filter's bandwidth that determine the resolution of adjacnt signals that are equal in amplitude.

In addition to bandwidth, there are other factors that determine useful resolution, such as filter type, filter shape (selectivity), residual FM and noise sidebands.



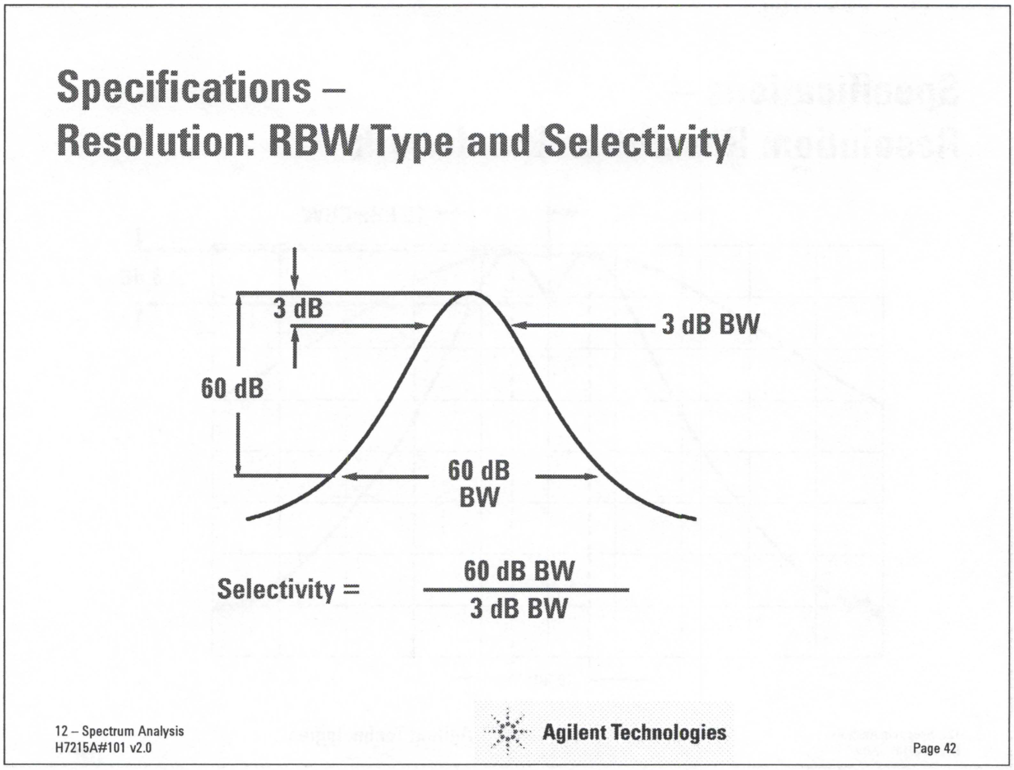
By adjusting the RBW to a smaller value, signals that are adjacent to each other may be resolved, The RBW is defined as either the 3dB or 6dB bandwidth of the filter. Agilent uses the 3dB definition.



When measuring two signals of equal-amplitude, the value of the selected RBW tells us how close together they can be and still be distinguishable from one another.

For example, if two signals are 10 kHz apart, a 10 kHz RBW will just separate the responses. A wider RBW may make the two signals appear as one.

In general, two equal-amplitude signals can just be resolved if their separation is equal to the 3 dB bandwidth of the selected resolution bandwidth filter.A better representation of two separate signals can be made if the RBW is less than the signal separation.



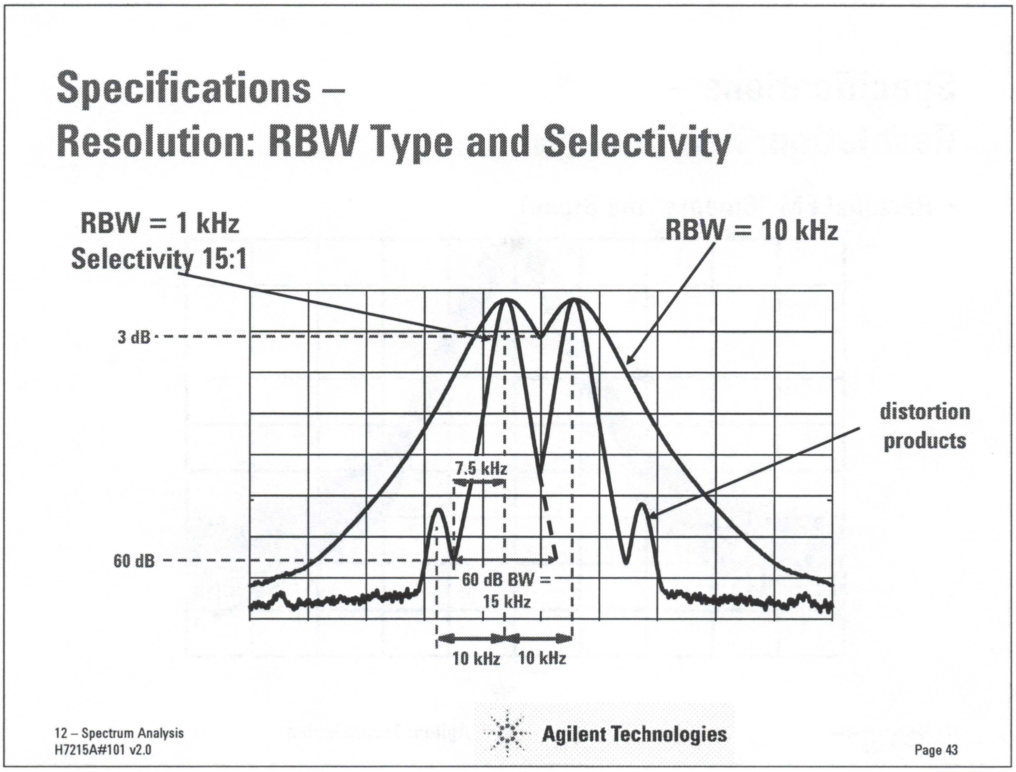
Selectivity (Shape Factor)

Selectivity is the ratio of the 60 dB to 3 dB (or 6dB) filter bandwidth. Typical selectivity's range from 11:1 to 15:1 for analog filters, and 5:1 for digital filters, in Agilent analyzers. If the RBW is just small enough to resolve signals at the same level, how well will unequal amplitude signals be shown? The greater the amplitude difference, the more a lower signal gets masked by the skirt of the adjacent response. Many close-in signals are distortion or modulation products and, by nature, are quite different in amplitude from the carrier signal.

Why such a large shape factor?

The spectrum analyzer is a receiver that has a mission to measure complex signals over wide frequency spans. This means sweeping as fast as possible while measuring as accuratly as possible. To measure at a fixed frequency a flat-top, low shape factor, fiter is good (e.g. a wave analyzer) but if a signal is swept through such a filter, considerable ringing occurs, or if a wide band (impulsive) signal is selected by such a filter the response is distorted. The theoretically best shape is a gaussian* shape. Such a filter is unrealizeable, but the "bell" shaped asynchronously tuned filter is a good compromise for the analog filters used in a spectrum analyzer.

* The convolution of an impulse and gaussian response gives a gaussian response.



Two tone intermodulation example

This is a two-tone test where the signals are separated by 10 kHz. The third-order distortion products are 10 kHz away from the two test tones and are 50 dB down.

With a 10 kHz RBW, the two equal-amplitude signals may be distinguished, but the distortion products are masked.

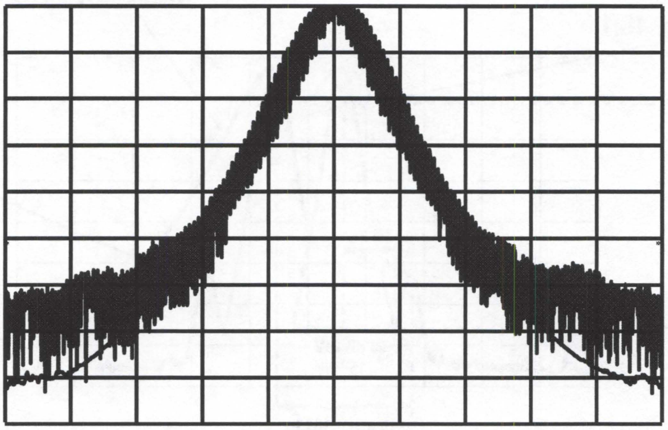
With a 3 kHz RBW with selectivity of 15:1. The filter width 60 dB down will therefore be 45 kHz (3x15). Still, the distortion products will be covered.

With a narrower filter, for example 1 kHz, the 60 dB bandwidth will be (1x15) 15 kHz. The distortion products are now visible.

In general, two signals that are unequal in amplitude by 60 dB must be separated by at least one half the 60 dB bandwidth in order to resolve the smaller signal.

Specifications – Resolution: Residual FM

- **Residual FM "Smears" the Signal**



12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 44

Another factor affecting resolution is the frequency stability of the spectrum analyzer's local oscillator. This is referred to as residual FM which is a short term stability measure of the LO.

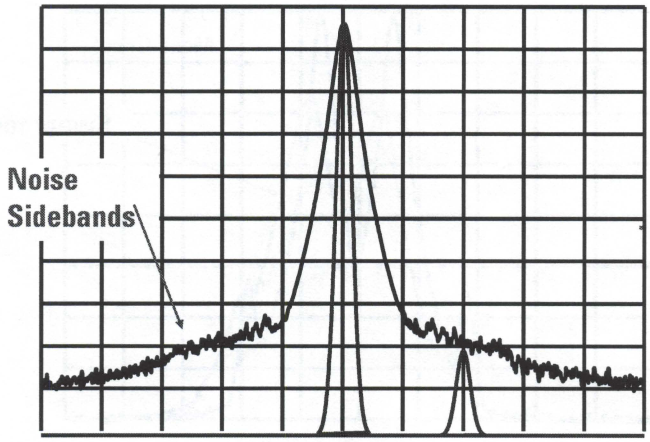
If the spectrum analyzer's RBW is less than the peak-to-peak FM, then this residual FM can be seen and looks as if the signal has been "smeared". This is because each sweep of the analyzer intercepts the signal which is moving, and shows as a time domain effect. This means that the spectrum analyzer's residual FM dictates the minimum RBW allowable, which in turn determines the minimum spacing of equal amplitude signals.

Phase locking the LOs to a reference reduces the residual FM and reduces the minimum allowable RBW. Higher performance spectrum analyzers are more expensive because they have better phase locking schemes with lower residual FM and smaller minimum RBWs.

Noise sidebands are also due to the short term stability, this time the RBW is much larger than any peak deviation, and show up as a frequency domain effect, i.e. as spectral density.

Specifications – Resolution: Noise Sidebands

- **Noise Sidebands can prevent resolution of unequal signals**



12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 45

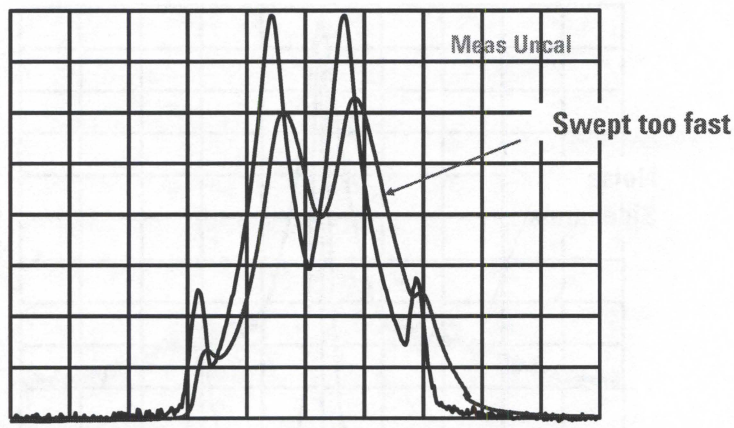
Noise Sidebands

Noise sidebands (also called phase noise) at the base of the signal response is the last factor in resolution, and can mask close-in (to a carrier), low-level signals. Phase noise is one of the limitations in an analyzer's ability to resolve signals of unequal amplitude. The above figure shows us that although we may have determined that we should be able to resolve two signals based on the 3-dB bandwidth and selectivity, we find that the phase noise masks the smaller signal. Noise sideband specifications are typically normalized to 1 Hz bandwidth. A phase noise specification of –80 dBc/1Hz at 10 kHz offset means that the noise trace at 10kHz from the carrier will be at least –50dBc if the RBW = 1 kHz due to the internal LO's . (1kHz = 1000 Hz will pass 30dB more noise power than 1Hz)

Be careful with terminology because the spectrum analyzer cannot distinguish between sideband spectral density due to phase or amplitude modulation, the term phase noise is technically incorrect. But because the mechanism that generates the noise sidebands is usually phase modulation, the term phase noise is often used. The specifications are important because there is no way to determine if displayed noise sidebands are due to the LO or the source being measured, except by knowing the specification.

Specifications –
Resolution: RBW Determines Measurement Time

- **Penalty For Sweeping Too Fast Is An Uncalibrated Display**



12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 46

Sweeptime

When the resolution bandwidth is decreased, the selected filter time constant increases and so it needs more time to fully respond. As sweeptime is decreased, for a given RBW filter, the displayed response will become progressivly distorted both in amplitude and frequency. The displayed amplitude becomes lower as the filter does not fully charge, and the frequency shifts higher due to delay through the filter.

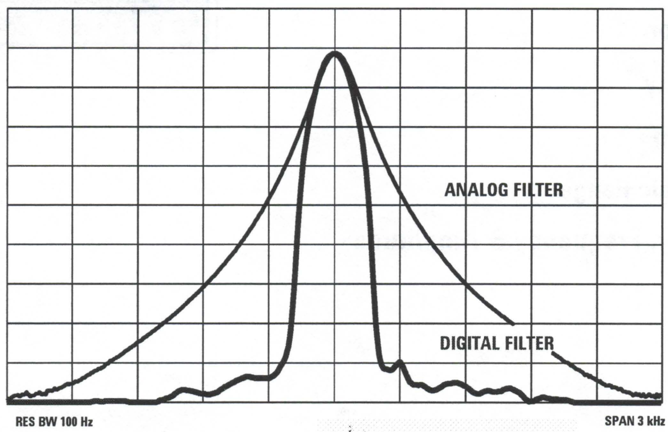
When selecting the RBW, there is usually a 1-10 or a 1-3-10 sequence of RBWs available (some spectrum analyzers even have 10% steps).
More RBWs are better because this allows choosing just enough resolution to make the measurement at the fastest possible sweeptime.

The spectrum analyzer firmware is programmed to adjust the sweeptime appropriatly for the RBW, VBW and SPAN selected, this happens automatically when the instrument is in its default (preset) mode. This is called the auto_coupled state.

$$\text{sweeptime is a function of } \left(\frac{Span}{RBW^2}\right) \text{ and } \frac{1}{VBW}$$

Specifications –
Resolution: Digital Resolution Bandwidths

- **Typical Selectivity**
  - **Analog 15:1 Digital 5:1**



Digital IF Filters

One thing to note before we close the topic of resolution is that Digital RBWs (that is, spectrum analyzers using digital signal processing (DSP) based IF filters) have superior selectivity and measurement speed compared to analog filters.

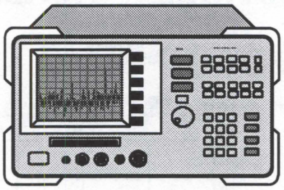
The table in your handout illustrates this point.

For example, with a 100 Hz RBW, a digital filter is 3.1 times faster than an analog.

RBW	Speed Improvement
100Hz	3.1
30Hz	14.4
10Hz	52.4
3Hz	118
1Hz	84

Specifications

- Safe Power Levels etc.
- Frequency Range
- Resolution
- Sensitivity
- Distortion
  - Dynamic Range
- Accuracy, Frequency & Amplitude



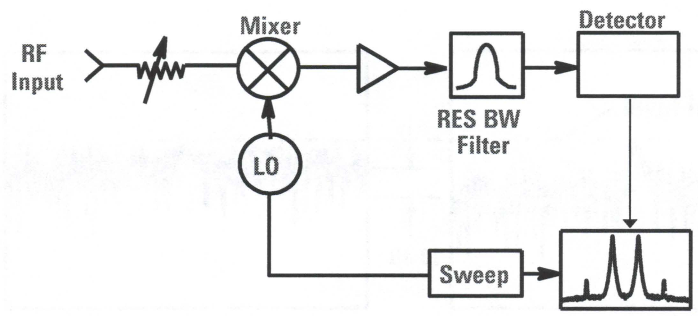
12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 48

Specifications – Sensitivity/DANL

- **A Spectrum Analyzer Generates and Amplifies Noise Just Like Any Active Circuit**



Sensitivity

The sensitivity of any receiver is an indication of how well it can measure small signals. The measurement of very low level signals is limited by the displayed noise level. A perfect receiver would add no additional noise to the thermal noise present at the input . All receivers, including spectrum analyzers, add noise, and this is generated predominantly in the first gain stage.

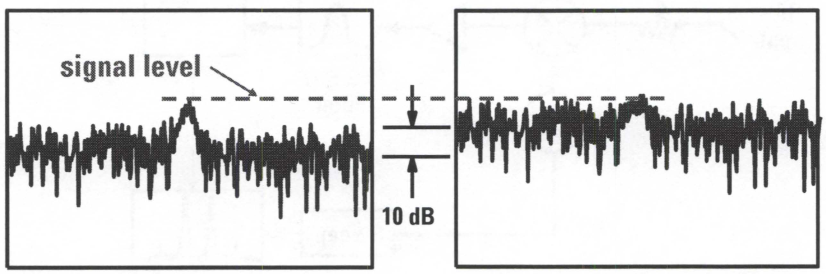
The displayed average noise level (DANL) in dBm, at a given RBW and input attenuator setting is the specification that determines sensitivity. It is the ultimate limitation in making measurements on small signals. An input signal below this noise level cannot be detected. Generally, sensitivity is on the order of -90 dBm to -145 dBm.

It is important to know the sensitivity capability of your analyzer in order to determine if it will adequately measure your low-level signals.

Note: In analyzers with an envelope detector, the DANL is technically different fron the actual noise level (the power spectral density) by a few dB due to offsets between the processing/detection of noise compared to a CW signal. Add 2.5dB to the displayed noise level to get power spectral density or use the noise marker.

Specifications - Sensitivity/DANL

- **Effective Level of Displayed Noise is a Function of RF Input Attenuation**
- **Signal-To-Noise Ratio Decreases as RF Input Attenuation is Increased**



Attenuation = 10 dB **Attenuation = 20 dB**

12 – Spectrum Analysis
H7215A#101 v2.0
Agilent Technologies
Page 50

One aspect of the analyzer's internal noise that is often overlooked is its effective level as a function of the RF input attenuator setting.

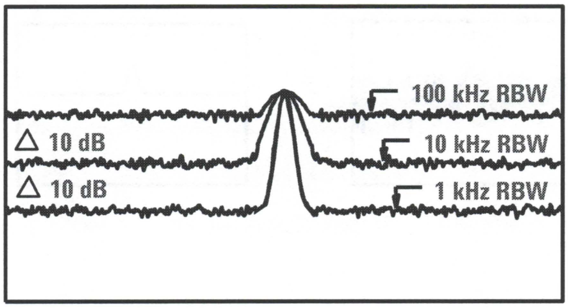
Since the internal noise is generated after the mixer, the RF input attenuator has no effect on the actual noise level. (Refer to the block diagram).

However, the RF input attenuator does affect the signal level at the mixer and therefore decreases the signal-to-noise ratio (SNR) of the analyzer.
The best SNR is with the lowest possible RF input attenuation.

Note in the figure, that the displayed signal level does not fall with increased attenuation. The RF input attenuator and IF gain are tied together so that the reference level will not change with attenuation. As the RF input attenuation is increased by 10 dB (–10dB), the IF gain increases 10 dB (+10dB). The result is that an on-screen CW signal stays constant, but the noise level is amplified and increases 10 dB.

Specifications –
Sensitivity/DANL: IF Filter (RBW)

- **Displayed Noise is a Function of IF Filter Bandwidth**
- **Decreased BW = Decreased Noise**



12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 51

The DANL is also a function of RBW

The internally generated noise in a spectrum analyzer is random and has no discrete spectral components. Also, its level is flat over a frequency range that is wide compared to the RBW ranges.

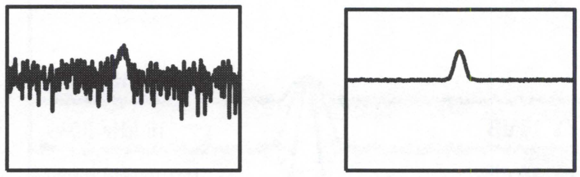
This means that the total noise reaching the detector (and displayed) is related to the RBW selected.

Since the noise is random, it is added on a power basis, so the relationship between displayed noise level and RBW is a ten log basis.

In other words, if the RBW is increased (or decreased) by a factor of ten, ten times more (or less) noise energy hits the detector and the displayed average noise level (DANL) increases (or decreases) by 10 dB.

Specifications – Sensitivity/DANL: VBW

- **Video BW Smoothes Noise for Easier Identification of Low Level Signals**



12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

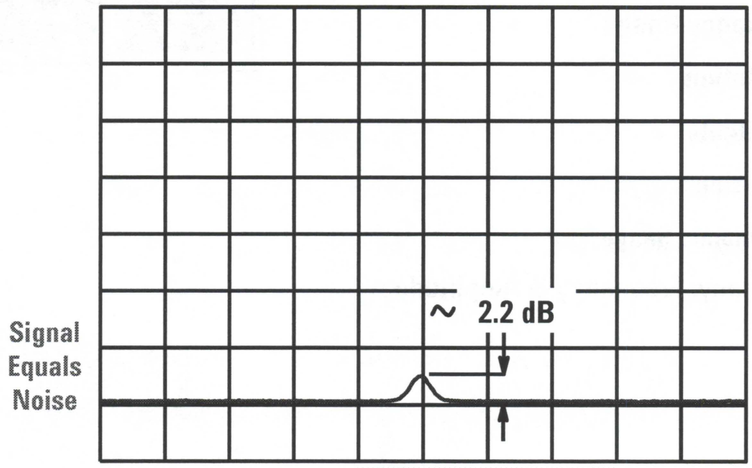
Page 52

The video filter(VBW) is used to smooth noise for easier identification of low level signals. The VBW, however, does not effect the frequency resolution of the analyzer (as does the RBW), and therefore changing the VBW does not improve sensitivity. It does, however, improve discern ability and repeatability of low signal-to-noise ratio measurements.

Note: In analyzers where the default video sampling is positive peak detection a decrease of VBW may appear to reduce the noise level. The average noise level is not changed, but the +ve peak detection was positively biasing the displayed noise.

Specifications – Sensitivity/DANL

- **Sensitivity is the Smallest Signal That Can Be Measured**



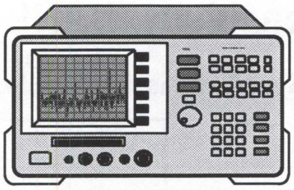
Sensitivity is then the smallest signal that can be measured

A signal whose level is equal to the displayed average noise level (DANL) will appear approximately as a 2.2 dB bump above the displayed average noise level.

This is considered to be the minimum measurable signal level.

Spectrum analyzer sensitivity is specified as the DANL in a specified RBW.

Specifications



- Safe Power Levels etc.
- Frequency Range
- Resolution
- Sensitivity
- Distortion
  - Dynamic Range
- Accuracy, Frequency & Amplitude

12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 54

Specifications – Sensitivity/DANL

- **For Best Sensitivity Use:**
  - **Narrowest Resolution BW**
  - **Minimum RF Input Attenuation**
  - **Sufficient Video Filtering (Video BW < .01 Res BW)**

Review

The best sensitivity is achieved with:

1. Narrowest RBW (decreases noise)
2. Minimum RF Input Attenuation (increases signal)
3. Using sufficient Video Filtering (to be able to see and read the small signal) (VBW less than or equal to 0.1 to 0.01 RBW)

Note however, that best sensitivity may conflict with other measurement requirements.

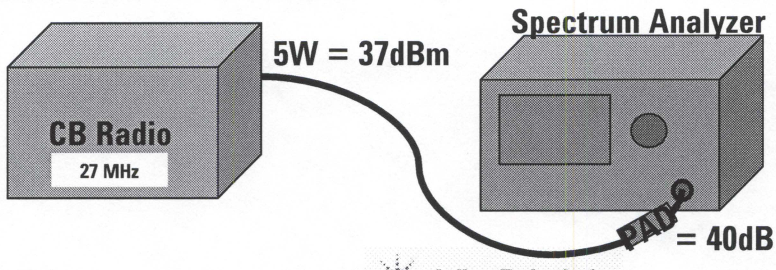
For example, smaller RBWs greatly increase measurement time.

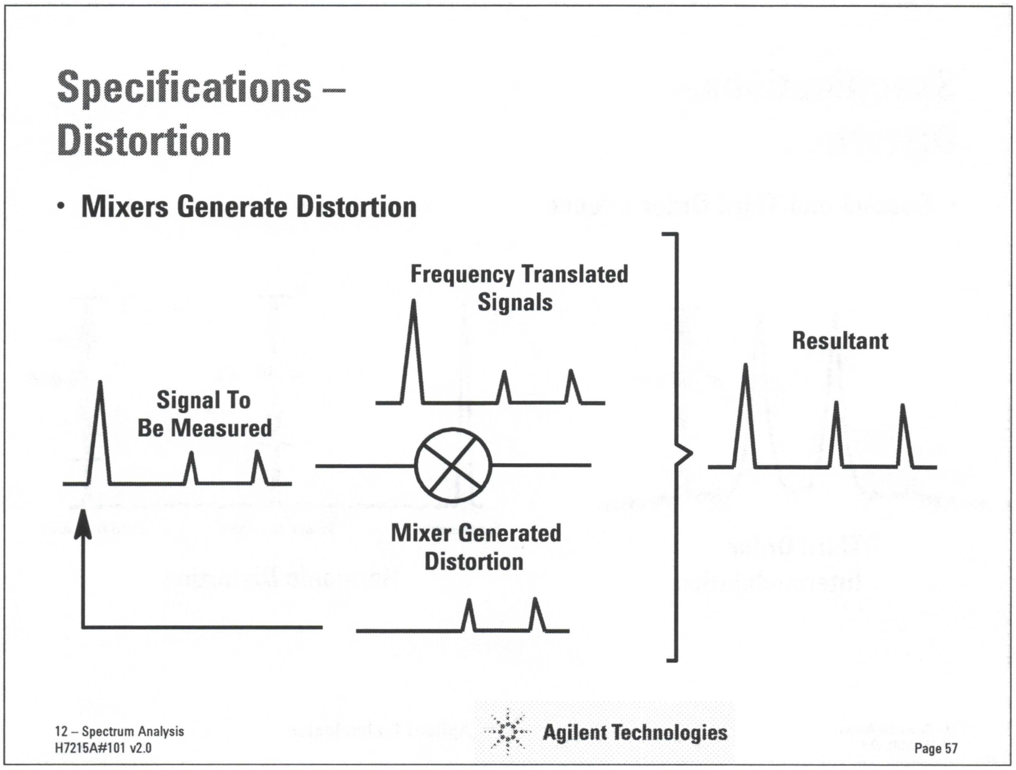
Also, zero dB input attenuation increases mismatch uncertainty therefore decreasing measurement accuracy.

Example of Dynamic Range

- What is the level of the second harmonic at 54MHz?
- Is the displayed signal at 54MHz coming from the radio?
- How do we make the measurement?

FCC regualtion:
second harmonic < –60dBc



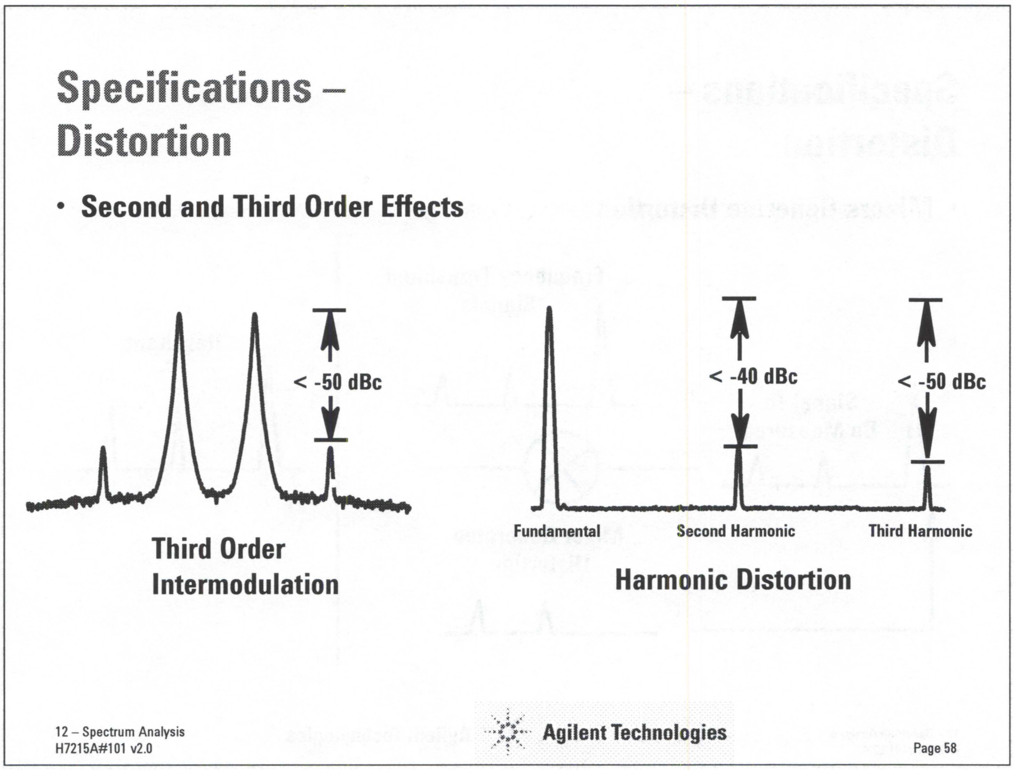


Distortion

Distortion measurements, such as 3rd order intermodulation and harmonic distortion, are common measurements for characterizing devices, the spectrum analyzer itself will also produce distortion products, and potentially disturb the measurement.

The distortion performance of the analyzer is specified by the manufacturer, either directly or lumped into a dynamic range specification.

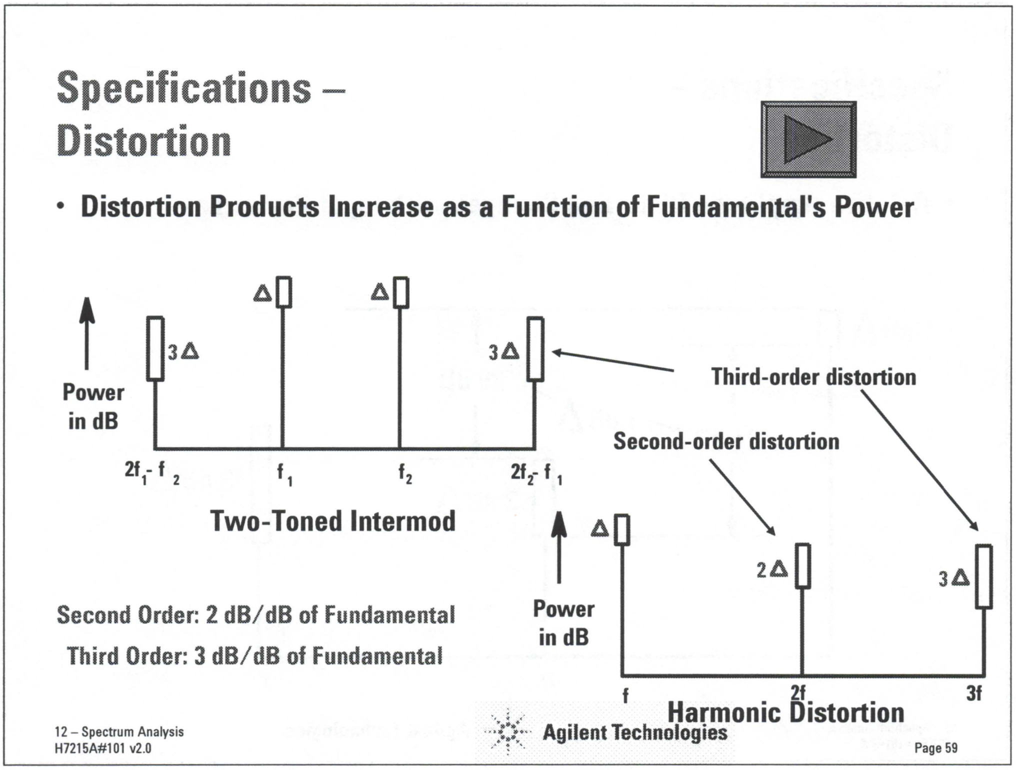
Because mixers themselves are non-linear devices, they will generate internal distortion. The response displayed is the vector sum of the desired signal and the internally generated signal, their relative phase being unknown, the measured signal will be subject to a large uncertainty, and, at worst, the internal distortion will completely mask the distortion products of the device to be measured.



These internally generated distortion products are a function of the amount of power present at the mixer, which is the input power minus the value of input attenuation. The measurement strategy is therefore to make sure that the internally generated products are below the expected products to be measured.

If the test specifies that the two-tone distortion products (third order products) must be more than, 50 dB below, and second order (harmonic) distortion more than 40 dB below, the fundamental, then this sets the minimum levels necessary for the analyzer specifications.

However, to reduce measurement error caused by the presence of internal distortion, the internal distortion must actually be much lower than the test specifications. To keep the uncertainty of measurement less than 1dB the Internal distortion products should be about 20dB below the expected measurement.



Any distortion products will increase as a function of the fundamental signal's power.

The behavior of distortion for any nonlinear device, whether it be the internally generated distortion of the spectrum analyzer's first mixer or the distortion generated by your device under test is shown in the slide.

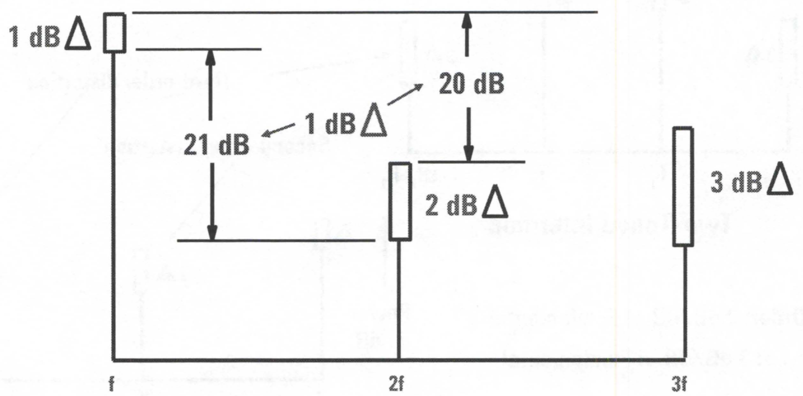
For example when the fundamental power is increased by 1dB, the second-order distortion will increase 2dB, and the third-order distortion will increase 3dB.

.

This means that on the log scale of our spectrum analyzer, the level of the second-order distortion will change twice as fast as the fundamental, and the third-order distortion will change three times as fast.

Specifications – Distortion

- **Relative Amplitude Distortion Changes with Input Power Level**



12 – Spectrum Analysis
H7215A#101 v2.0
Agilent Technologies
Page 60

Since most distortion measurements are made relative to the fundamental signals (the carrier or two-tones), let's re-examine the behavior in relative terms.

When the fundamental power is changed 1 dB, the second-order distortion changes absolutely by 2 dB, but relative to the fundamental, the second-order distortion changes 1 dB. (1dBc)

There is a one-to-one relative relationship between the fundamental and second-order distortion.

When the fundamental power is changed 1 dB, the third-order distortion changes absolute by 3 dB, but relative to the fundamental, the third-order distortion changes 2 dB. (2dBc)

There is a two-to-one relative relationship between the fundamental and third-order distortion.

The reason for the 2× and 3× change in absolute level may be justified by considering the application of a sinusoidal small signal to a non-linear characteristic which can be described by a polynomeal $y = a_1x + a_2x^2 + a_3x^3 + \ldots$

Let $x = m.\cos \omega t$.

Since $\cos^2 A = (\cos 2A + 1)/2$ *and* $\cos^3 A = (\cos 3A + 3\cos A)/4$

Then terms like $a^2m^2/2 \cos 2\omega t$ and $a^3m^3/4 \cos 3\omega t$ are formed; m is the magnitude term which when increased by a small quantity "h" and expanded becomes $(m+h)^2 = m^2 + h^2 + 2mh$ for the $\cos 2\omega t$ term and $m^3 + h^3 + 3m^2h + 3mh^2$ for the $\cos 3\omega t$ term. If the terms in h^2 and h^3 are considered negligable then:

The ratio of the increase for the

fundamental: $(m + h)/m = 1 + h/m$

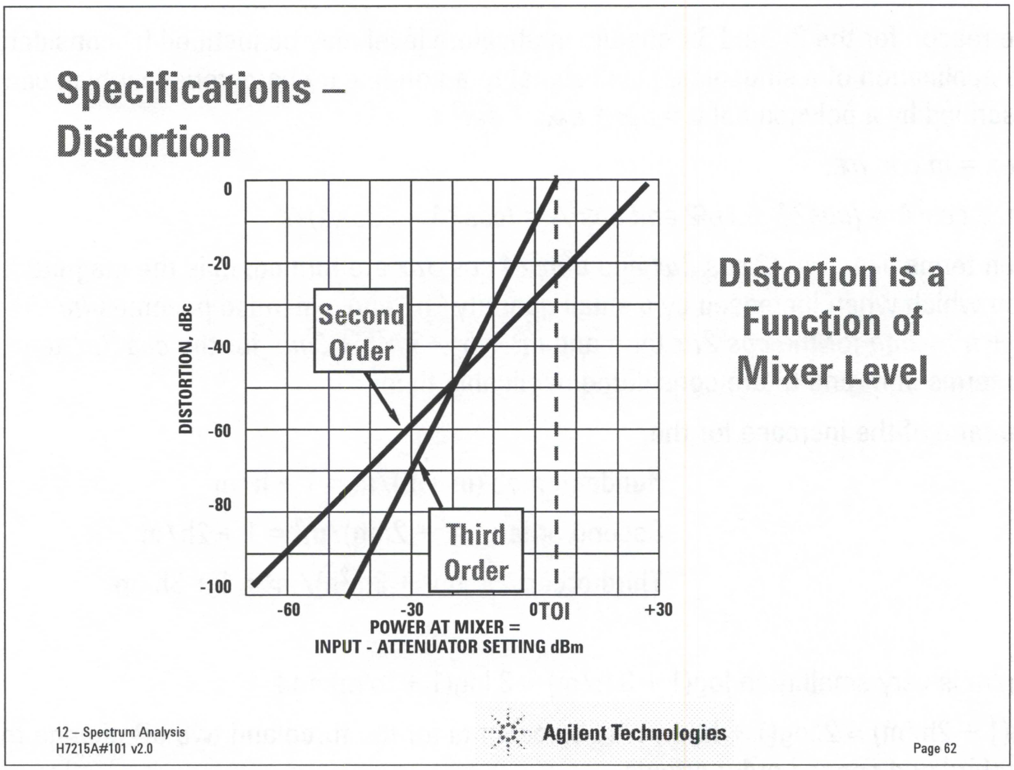
Second order: $(m^2 + 2mh)/m^2 = 1 + 2h/m$

Third order: $(m^3 + 3m^2h)/m^3 = 1 + 3h/m$

If h/m is very small then $\log(1 + 3\,h/m) \approx 3.\log(1 + h/m)$ and

$\log(1 + 2h/m) \approx 2.\log(1 + h/m)$ which accounts for the three and two dB change for the third and second order effects.

Note: 1dB is about 26% in power which is quite a large change.



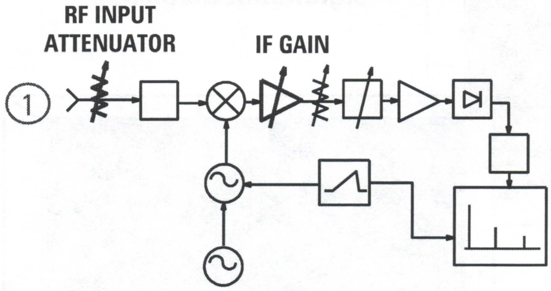
Understanding the relative relationship is useful in determining distortion within the analyzer. Here is a graph of the level of the second and third order distortion products relative to the signals that cause them:

- •The x-axis is the signal power at the first mixer. The y-axis is the spectrum analyzer's internally- generated distortion level in dBc. These curves are called signal-to-distortion curves.
- •The slope is unity for the second-order, because remember there is a 1-for-1 relative relationship.
- •The third-order curve has a slope of two - there is a 2-for-1 relative relationship.
- •Therefore, if analyzer distortion is specified for one signal level at the mixer, distortion at any other level can easily be determined.

This example shows that for a level of –40 dBm at the mixer, third-order distortion is –90 dBc and second-order distortion is –70 dBc. The mixer level at which 3rd-order distortion equals the fundamental (0 dBc) is useful to know. It gives us a simple expression that permits computation of 3rd-order distortion at any mixer level. This reference point is called the third-order intercept or TOI and is a condition which could never happen because compression in the mixer would occur first. TOI is a common spectrum analyzer specification, and is used to determine the maximum dynamic range available for a particular measurement. In the above figure, TOI = +5 dBm.

Specifications – Distortion

- **Distortion Test:**
  - **Is it Internally or Externally Generated?**



Change Input Attn by 10 dB

(2) Watch Signal on Screen:

→ No change in amplitude = distortion is part of input signal (external)

→ Change in amplitude = at least some of the distortion is being generated inside the analyzer (internal)

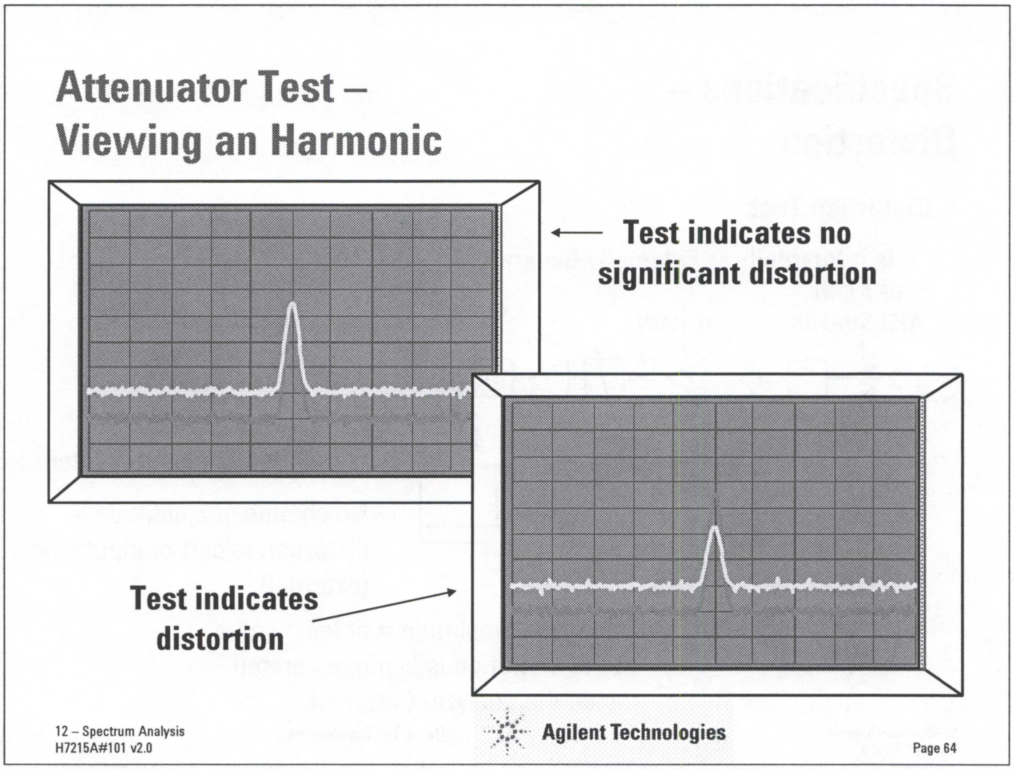
12 – Spectrum Analysis
H7215A#101 v2.0
Agilent Technologies
Page 63

Before leaving this section on distortion, there is a test that should be done for all distortion measurements which will tell us whether or not what we are seeing on the screen is internally generated distortion, or distortion caused by the DUT.

Remember that the RF input attenuator and the IF gain are tied together such that input signals will remain stationary on the screen when we adjust the RF input attenuation. So let's change the RF input attenuation and see what happens.

If the distortion product on the screen does not change, we can be sure it is distortion from the DUT (i.e. part of the input signal). The 10 dB attenuation applied to the signal is also experiencing the 10 dB gain from the IF gain and therefore, there is no change.

If however, the signal on the screen does change, then we know it must be being generated, at least in part, somewhere after the input attenuator, and not totally from the DUT. The 10 dB attenuation is not applied to this internal signal (since it is actually generated after the attenuator), yet the 10 dB gain is applied to it, therefore increasing its level by as much as 10 dB.



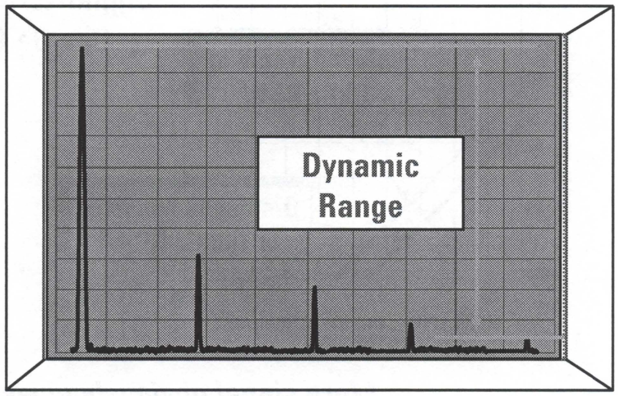
This could be the view of a third harmonic, in one case applying 10dB internal attenuation has no effect on the amplitude of the signal, this would indicate a good measurement. In the second case the amplitude has decreased. The theory would indicate that, if there was internal distortion only causing this response, then the change in amplitude should be 30dB. This display indicates that there is a combination of external and internal third harmonic.

It is important to know that just because we may determine that there is no measurement distortion for this particular signal, other signals at other frequencies may be subject to measurement distortion. Non linear effects are not well behaved. After one step of the attenuator another step, a further step may show no change, so measurement can be made at the first attenuator position. Because setting an optimum signal level at the mixer is the key to achieving the largest dynamic range, there is a trend to providing spectrum analyzers with RF attenuators with 5dB steps.

Using an external attenuator

An external attenuator can be used to get the best control of signal level. Remember the IF amplifier gain will not change, the amplitude of the displayed signal will change by the amount of attenuation. In this case watch for *relative change* between the fundamental and the suspected high order component.

Specifications – Dynamic Range



Dynamic Range is defined as the maximum ratio of two signal levels simultaneously present at the input which can be measured to a specified accuracy.

Imagine connecting two signals to the analyzer input - one which is the maximum allowable level for the analyzer's input range and the other which is much smaller.

The smaller one is reduced in amplitude until it is no longer detectable.

When the smaller signal is just measurable, the ratio of the two signal levels (in dB) defines the dynamic range of the analyzer.

What effects might make it undetectable?

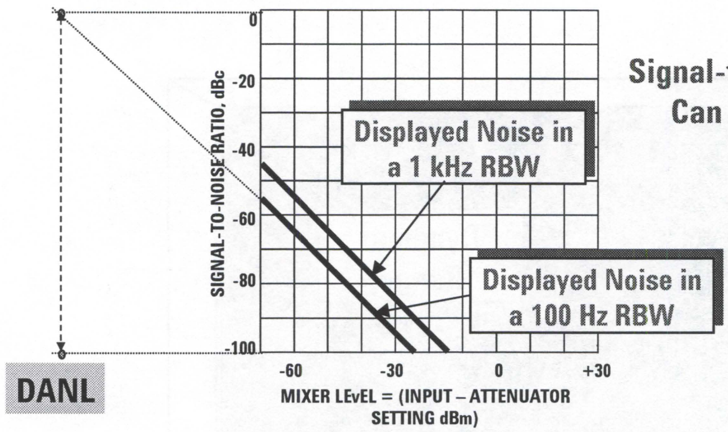
•Such things as residual responses, distortion, and the internal noise of the analyzer. These will all be large enough to cover up the smaller signal as we decrease its amplitude

.

The dynamic range of the instrument determines the amplitude range over which we can reliably make measurements.

Specifications – Dynamic Range

Signal-to-Noise Ratio Can Be Graphed



More signal obviously gives more S/N

12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

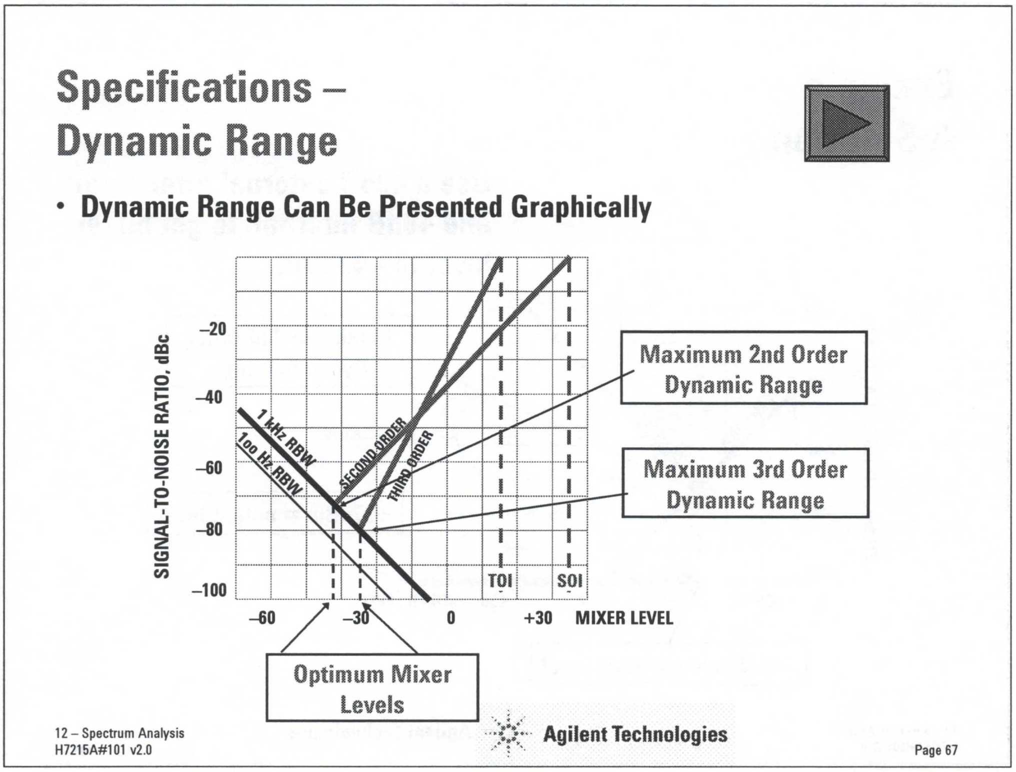
Page 66

We can also plot signal-to-noise ratio (SNR) as a function of mixer power, on the dynamic range graph. A family of curves can be plotted by knowing the DANL at various RBW's , the lines slope down to the right at a one to one slope. More signal obviously gives more S/N , that is a larger dBc number.

The previously plotted signal-to-distortion curves showed the dynamic range for distortion (minimum distortion in dBc) as a function of power level to the input mixer, the lowest power to give required dynamic range plus margin should be chosen. These SNR curves, however, tell us that best dynamic range for noise occurs at the highest signal level possible.

This a classic engineering compromise, on the one hand, we would like to drive the level at the mixer to be as large as possible for the best signal-to-noise ratio. But on the other hand, to minimize internally generated distortion, we need as low a drive level to the mixer as possible.

The best dynamic range is a compromise between signal-to-noise and internally generated distortion.



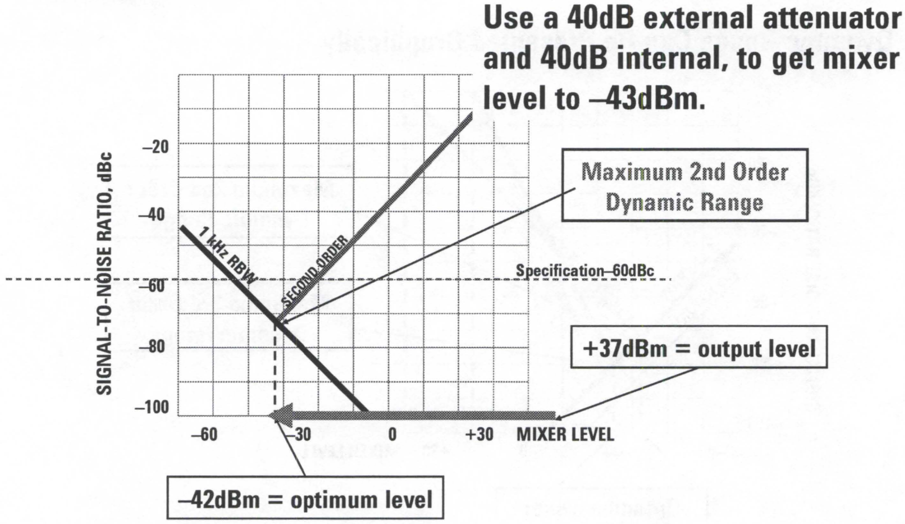
A graph of both the signal-to-noise and signal-to-distortion curves are here plotted together.

Optimum dynamic range is where the curves intersect, where the internally generated distortion level equals the noise level. This shows the dynamic range specifications for 2nd-order and 3rd-order distortions.

If the test tones are 0 dBm and the attenuator has 10 dB steps. Mixer levels of 0, –10, – 20, – 30, – 40 dBm, etc. may be set. Several settings will give enough dynamic range to measure 3rd-order distortion products at – 60 dBc. The internal noise and distortion products must be as low as possible to minimize errors. A drive level at the mixer between – 30 and – 40 dBm would allow us to make the measurement with minimum error. Which mixer level do we choose?

For < 1 dB uncertainty in your measurement, the signal-to-internal-distortion must be 19 dB, whereas the signal-to-noise only 5 dB. This tells us that it is best to stay closer to the noise, so we would set mixer level to – 40 dBm (the mixer level to the left of the third-order point of intersection). This results in a "spurious free display".

Example – A Solution



12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 68

The total attenuation should be about 80dB to get the mixer level to within a level to ensure at least 10dB margin. 20dB is necessary for a < 1dB uncertainty. This could be done by using a 100Hz RBW and a further 10dB attenuation.

Specifications – Dynamic Range

- **Calculated Maximum Dynamic Range**

$$MDR_3 = 2/3\,(DANL - TOI)$$

$$MDR_2 = 1/2\,(DANL - SOI)$$

Where TOI = Mixer Level - dBc/2

SOI = Mixer Level - dBc

Optimum Mixer Level = DANL - MDR

Attenuation = Signal - Optimum Mixer Level

12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 69

We can calculate maximum dynamic range using these equations, where:

- MDR3 = maximum third-order dynamic range, and
- MDR2 = maximum second-order dynamic range

To calculate these, we use the DANL, TOI and SOI values typically given on the datasheet, where:

- TOI = Third-order intercept
- SOI = Second-order intercept
- DANL = Displayed average noise level

Once we've calculated MDR, we can determine the optimum mixer level:

- Mixer level = signal level – attenuation
- Optimum mixer level = mixer level for maximum dynamic range

Specifications – Dynamic Range

- **Example Calculation**

$MDR_3 = 2/3\ [(-115) - (+5)]$
$= -80$ dBc (1 kHz RBW)

Where TOI = (-30) - (-70)/2
= + 5 dBm

Optimum Mixer Level = (-115) - (-80)
= -35 dBm

Attenuation = (0) - (-35)
= +35 dBm

12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 70

Let's do an example

Let's say we have a spectrum analyzer with a DANL = -115 dBm (1 kHz RBW), and TOI = +5 dBm.

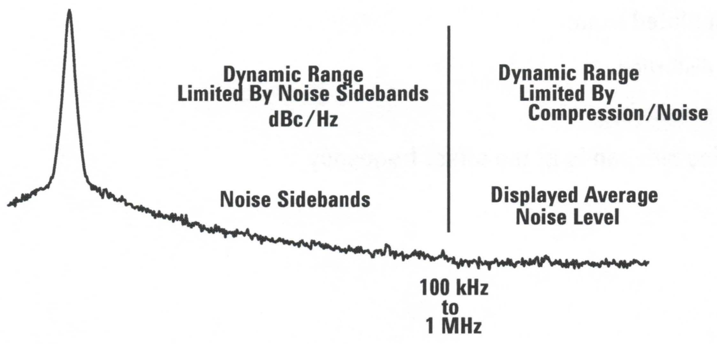
This slide shows how to calculate maximum third-order dynamic range (MDR3), optimum mixer level, and attenuation.

Remember that for every order of magnitude decrease in RBW, the DANL decreases by 10 dB.

Therefore, if we decrease our RBW to 10 Hz, DANL = -135, and third-order dynamic range improves by 13 dB, [2/3(-140)] = 93 dBc.

Specifications – Dynamic Range

- **Dynamic Range for Spur Search Depends on Closeness to Carrier**



12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 71

The final factor in dynamic range is the phase noise, or noise sidebands, on our spectrum analyzer LO.

An example application where we can see how both the noise sidebands and the DANL limits dynamic range is when making spur measurements. As shown on the slide, the dynamic range for the close-in, low-level spurs is determined by the noise sidebands within approximately 100 kHz to 1 MHz of the carrier (depending on carrier frequency). Beyond the noise sidebands, the dynamic range is limited by DANL

Another example is when the signals are so close together that noise sidebands limit dynamic range (e.g. a two-tone measurement where the tones are separated by 10 kHz, therefore producing third-order distortion products 10 kHz from the test tones).

For distortion tests, the phase noise can also be plotted on the dynamic range graph as a horizontal line at the level of the phase noise specification at a given offset.

NOTE: The dynamic range curves we've just discussed are needed only for distortion tests.

Specifications – Dynamic Range

- **Actual Dynamic Range is the Minimum of:**
  - **Maximum dynamic range calculation**
  - **Calculated from:**
    - **distortion**
    - **sensitivity**
  - **Noise sidebands at the offset frequency**

12 – Spectrum Analysis
H7215A#101 v2.0

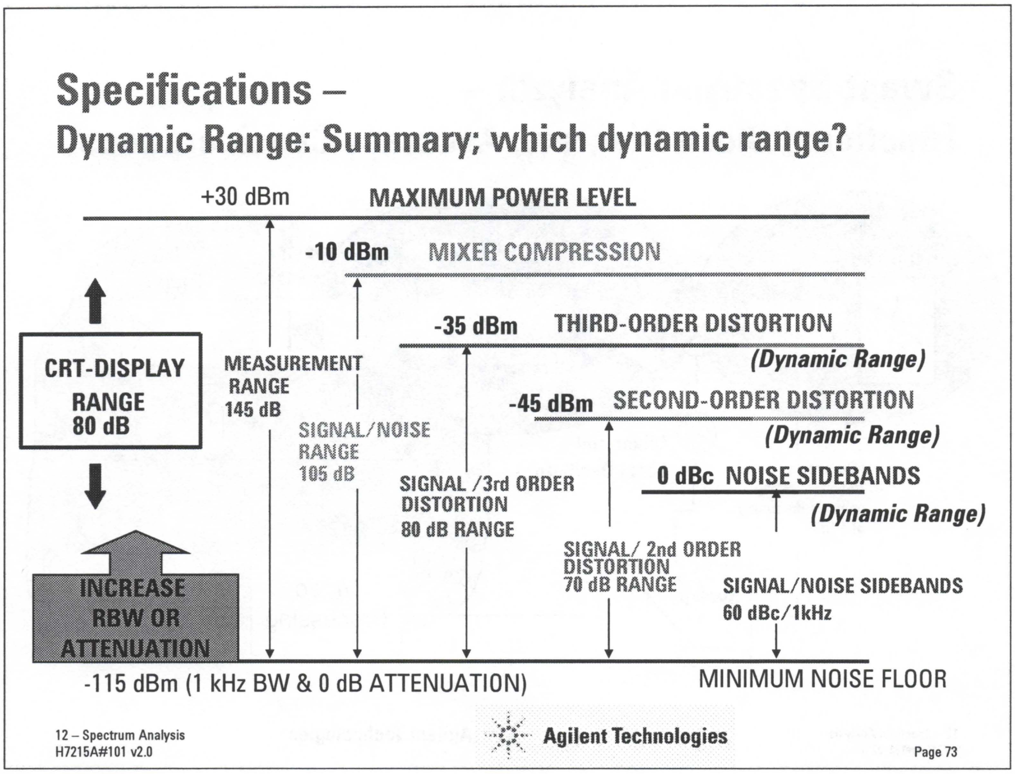
Agilent Technologies

Page 72

We have seen that the dynamic range of a spectrum analyzer is limited by three factors: the broadband noise floor (sensitivity) of the system, the distortion performance of the input mixer, and the phase noise of the local oscillator.

The first two factors are used to calculate maximum dynamic range.
Therefore, actual dynamic range is the minimum of:

- The MDR calculation and
- The noise sidebands.

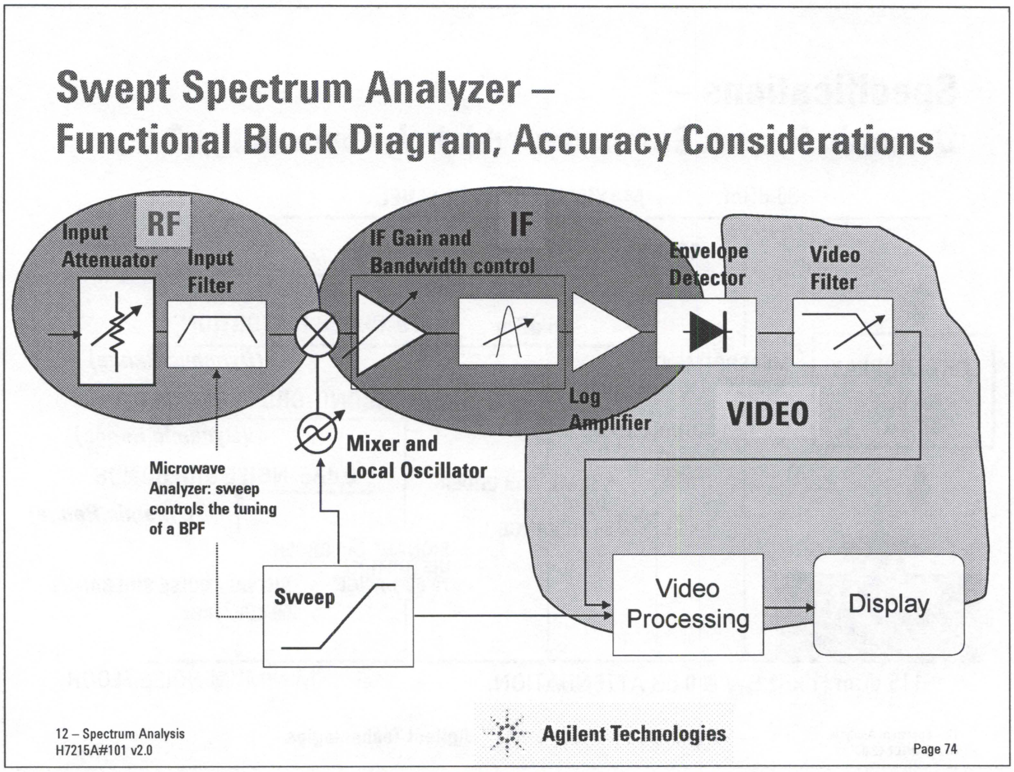


In summary, let's discuss all the "ranges" associated with a spectrum analyzer. Typically the term "dynamic range" only refers to the ability to measure two signals at the same time.

- **Display range** refers to the calibrated amplitude range of the CRT display. For example, some analyzers with a display having 8 divisions might only have a 70 dB display range (with 10 dB/ div) because the bottom division is not calibrated, while others will have a full 100dB display and calibrated amplitude range.
- **Measurement range** is the ratio of the largest to the smallest signal that can be measured under any circumstances - not at the same time. The upper limit is determined by the maximum safe input level, +30 dBm (1 Watt) for most analyzers. Sensitivity sets the other end of the range.
- **Dynamic Range:. . . *it depends on what you are measuring.***

The other four ranges (signal/noise, signal/third order distortion, signal/second order distortion, and signal/noise sidebands) are when measuring two signals at the same time, and therefore are called dynamic range specifications.

The ability of the analyzer to simultaneously large and small signals is limited by the dynamic range, When specifying dynamic range, it is important to qualify the term with the type of measurement to be made. The various specifications which can improve dynamic range will relate to the cost of the spectrum analyzer.



Block Diagram Accuracy Considerations

If the functional block diagram is divided into its signal flow stages, the business of amplitude accuracy is easier to handle.

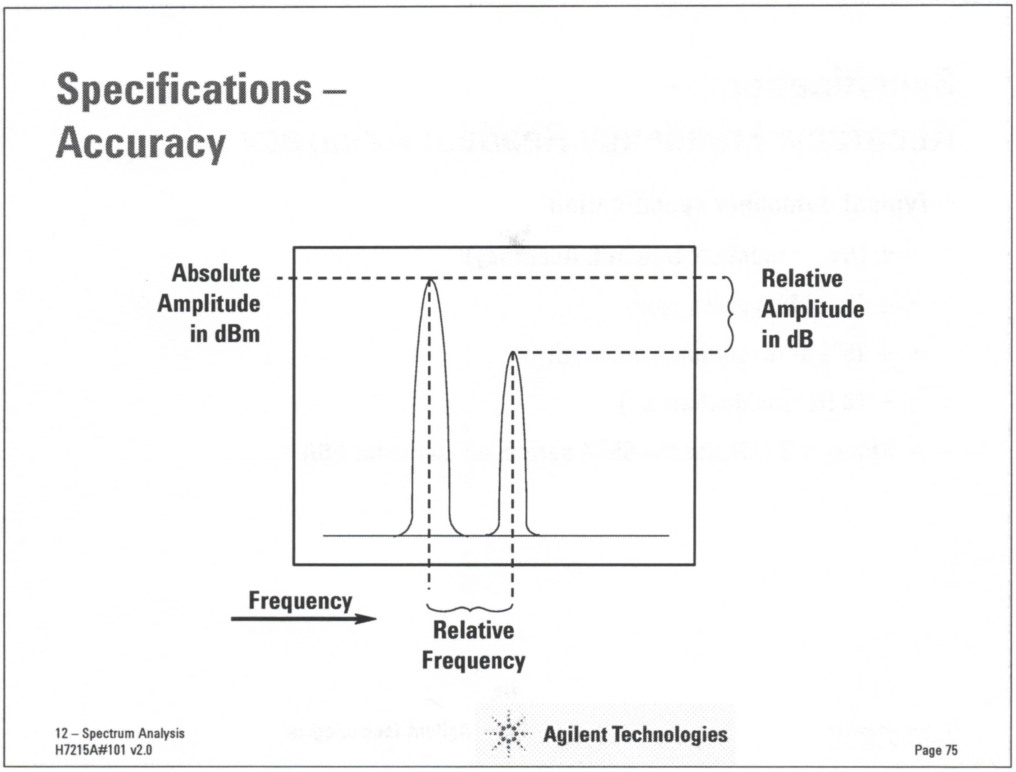
The RF section must deal with broadband signals, assuming linearity the major factors are:

- Frequency response of the filter, attenuator (at each step) and the conversion loss of the mixer as a function of frequency.
- For a microwave analyzer, the conversion loss must be known for different harmonics of the LO, also the pre-selector may have an insertion loss which is a dynamic function of frequency.

The IF section deals with a narrowband signal, no more than a few MHz, whereas if the RF section is calibrated at one RF frequency the IF section is well calibrated for any RF signal. The major factors are:

- IF filter changes
- IF gain changes
- Log amp law fidelity.

Although the log amplifier in in this section its contribution to measurement uncertainty is usually lumped in with display fidelity.



Accuracy

All measurements are estimates, and as such are subject to measurement uncertainty. The term *uncertainty* is preferred by metrologists however we will call it accuracy.

If we state a level is – 45dBm, there must be an implied accuracy. In conversation this may not be stated, however in formal communications the level should be written, for example, – 45dBm.±0.5dB. The specification of the spectrum analyzer will include statements of measurement accuracies.

Measurement data taken from a spectrum analyzer is both *absolute* and *relative,* for example frequency or power , and frequency or power, change or difference.

An absolute measurement may be made with a single marker. For example, the frequency and power level of a carrier.

A relative measurement may be made with a delta marker and displays the frequency or amplitude of one signal relative to another.

Specifications –
Accuracy: Frequency Readout Accuracy

- **Typical datasheet specification**
  - **± (freq. readout × freq. ref. Accuracy)**
  - **+ 1% of frequency span**
  - **+ 15% of resolution bandwidth**
  - **+ 10 Hz "residual error")**
  - **Spans < 2 MHz:for the 859X series; all spans for ESA**

12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 76

Frequency Accuracy

The most important frequency accuracy specification is 'Frequency Readout Accuracy.' There are several sources of uncertainty:

1) The Frequency Reference accuracy depends on the basic architecture of the analyzer. Most modern analyzers have a frequency stabilized LO frequency, this is referred to as "synthesized" although the technical means of achieving the stabilization may not be a classic synthesizer.
Some or all of the oscillators are referenced to a single, traceable, reference oscillator and have accuracy's of a few hundred hertz.

2) Span error is typically split, having a small span and a large span specification. This is because the stabilization method may change between large and small spans.

3) RBW error is typically much smaller than span error, this is an uncertainty due to the tolerance of the actual center frequency of RBW filter from the designed intermediate frequency.

4) There is also a small "residual error" to account for in frequency readout accuracy.

5) In microwave analyzers when harmonic mixing is used the "N" of the mixing product will figure in the calculation.

Specifications –
Accuracy: Frequency Readout Accuracy Example

- **Single Marker Example:**
  - **2 GHz**
  - **400 kHz span**
  - **3 kHz RBW**

Calculation:	$(2 \times 10^9\ \text{Hz}) \times (1.3 \times 10^{-7}/\text{year})$	=	**260Hz.**
	1% of 400kHz span	=	**4000Hz.**
	15% of 3kHz RBW	=	**450Hz.**
	10 Hz residual error	=	**10Hz.**
	Total	=	**± 4720Hz.**
			$\boxed{\pm 4.72\ \text{kHz.}}$

12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 77

Here is an example calculation to find the frequency accuracy at a particular measurement condition.

If measuring a signal at 2 GHz, using a 400 kHz span and a 3 kHz RBW, then

Frequency reference error = (ageing rate × period of time since adjustment + initial achievable accuracy + temperature stability)

So, we can say that our signal could actually be anywhere from 2.0000047 to 1.9999953 GHz.

Specifications – Accuracy: Relative Amplitude Accuracy

- Display fidelity
- Frequency response
- Δ RF Input attenuator
- Δ Reference level
- Δ Resolution bandwidth
- Δ CRT scaling

12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 78

Amplitude Accuracy

When making relative measurements, some part of the signal is used as a reference. For example, when making second-harmonic distortion tests, the fundamental is the reference.

The absolute value of the fundamental or the second harmonic do not affect the measurement - only how the second harmonic differs in amplitude from the fundamental is of interest.

Relative amplitude accuracy depends on several factors.

The first two on the list will directly affect the accuracy, whereas the bottom four will only affect accuracy when they are changed.

Specifications –
Accuracy: Relative Amplitude Accuracy - Display Fidelity

- **Applies when signals are not placed at the same reference amplitude**
- **Display fidelity includes**
  - **Log amplifier or linear fidelity**
  - **Detector linearity**
  - **Digitizing circuit linearity**
- **Technique for best accuracy**

Display Fidelity

This is an error that occurs when the two signals being measured are not placed at the same reference level.

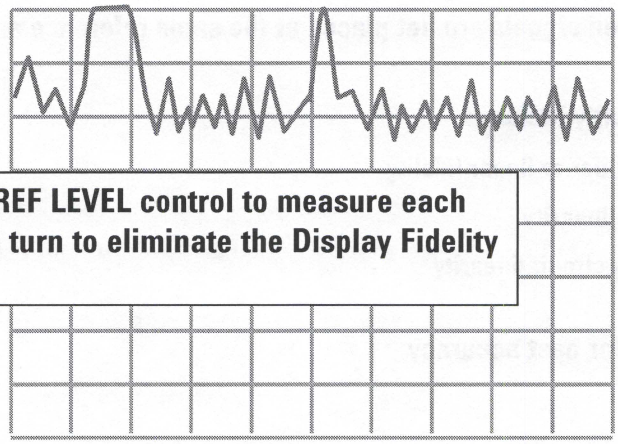
It covers a variety of factors such as: how true is the log characteristic of the log amplifier? how linear is the detector? and how linear is the digitizing circuitry?

The display fidelity is better over small amplitude differences, and ranges from a few tenths of a dB for signal levels close together to perhaps 2 dB for large amplitude differences.

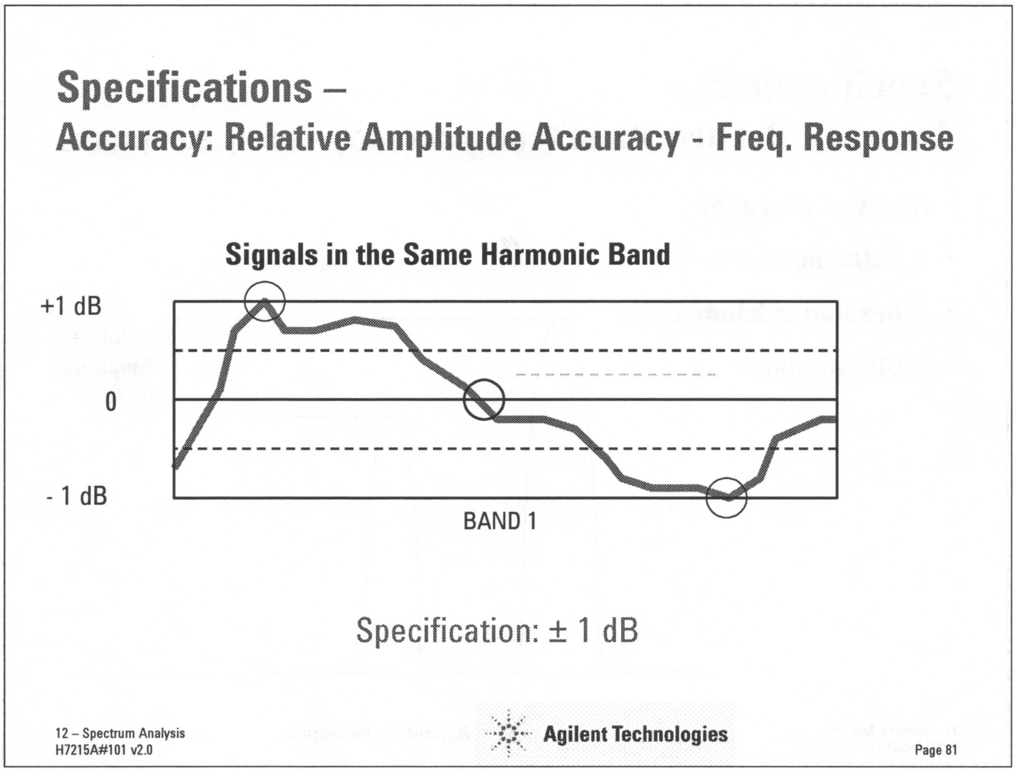
A technique for improving amplitude accuracy is to place the first signal at a reference amplitude using the reference level control, and use the marker to read amplitude value. Then move the second signal to the same reference and calculate the difference.

This assumes that the Reference Level Uncertainty (changing the reference level) is less than the Display Fidelity Uncertainty.

Measurement by IF Substitution



Use the REF LEVEL control to measure each Signal in turn to eliminate the Display Fidelity Error.



The second factor on our list of relative amplitude accuracy was frequency response.

The frequency response, or flatness, of the analyzer also affects relative amplitude accuracy, and is frequency range dependent.

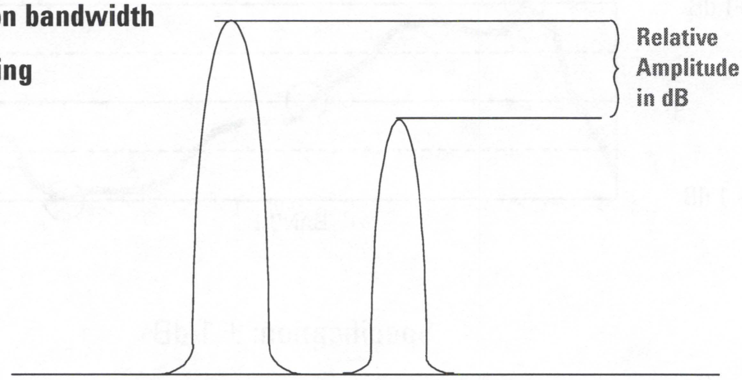
For example, a low-frequency RF analyzer might have a frequency response of ± 0.5 dB whereas a microwave analyzer in the 20 GHz range could have a frequency response in excess of ± 4 dB.

The specification assumes worst case, where the flatness varies the full amplitude range (as shown in the slide).

With the above specification of ± 1 dB, the uncertainty between two signals in this band would be 2x ± 1 dB = ± 2 dB, because we have no knowledge of the detail shown, and can only assume the full uncertainty applies to any signal in the range.(Metrologists call this a rectangular uncertainty).

Specifications –
Accuracy: Relative Amplitude Accuracy - Freq. Response

- **RF Input attenuator**
- **Δ Reference level**
- **Δ Resolution bandwidth**
- **Δ CRT scaling**



12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 82

The four other factors mentioned earlier, involve control changes made during the course of a measurement, and therefore can be eliminated if they can be left unchanged.

The RF input attenuator and IF gain (reference level control) have uncertainties associated with them. Since the RF input attenuator must operate over the entire frequency range of the analyzer, its accuracy, like frequency response, is a function of frequency. The IF gain only operates at the IF, and is therefore more accurate.

Since each filter will have a different insertion loss, a change the RBW between measurement steps can also degrade accuracy.

A change of display scaling between 10 dB/div to 1 dB/div, or between log and linear will also introduce uncertainty in the amplitude measurement.

Specifications – Accuracy: Absolute Amplitude Accuracy

- **Calibrator accuracy**
- **Frequency response**
- **Reference level uncertainty**

Absolute amplitude measurements are actually measurements that are relative to the calibrator, which is a signal of known amplitude.

Most modern spectrum analyzers have a built-in calibrator which provides a signal with a specified amplitude at a given frequency.

This calibrator source operates at one frequency, so we rely on the frequency response and other relative accuracies of the analyzer to extend absolute calibration to other frequencies and amplitudes.

A typical calibrator has an uncertainty of +/- 0.3 dB.

Since the input signal is probably at a different frequency to the calibrator and at a different amplitude, the absolute amplitude accuracy depends on calibrator accuracy, frequency response, and reference level uncertainty (also known as IF gain uncertainty) or the display fidelity, depending on the measurement method used.

Specifications – Accuracy: Other Sources of Uncertainty

- **Mismatch (RF input port not exactly 50 ohms)**
- **Compression due to overload (high-level input signal)**
-
- **Distortion products**
- **Signals near noise**
- **Microwave preselector repeatability**

12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 84

Other sources of Uncertainty

There are many other sources of uncertainty, some of these can be avoided, such as compression or distortion. Others such as signals near noise can be corrected.

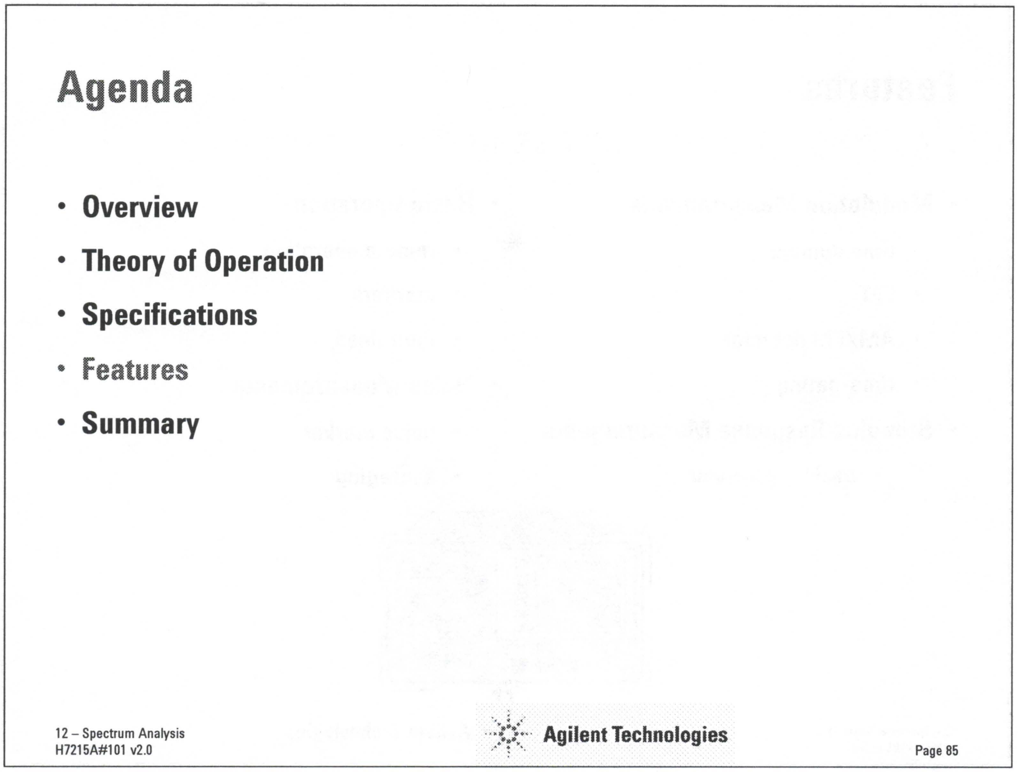
Worst Case Problem

When calculating uncertainty on a worst case basis the total number may be large ± many dBs. This is the case for microwave measurements in particular.

Example: A marker indicates -50dBm at 15GHz on an 8563E calibrated; 25°C;no change of RBW ;no change of Ref level; no change of RF attenuation.

Uncertainty	Frequency response, absolute/absolute typ.	± 4.0 dB/ ±2.5dB
	Scale Fidelity	± 0.85dB
	Calibrator(at 300MHz.)	± 0.3 dB
	TOTAL	± 5.15dB
	TOTAL using typical Frequency Response. ± 3.65dB	

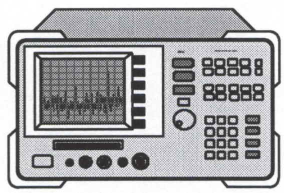
These high values can be a surprise, most system users will overcome this situation by doing calibrations at the frequencies/levels of the expected signal range. A calibration in software will create a lookup table to be referenced by the system as it makes measurements.



Both hardware and firmware features can make the spectrum analyzer much more useful. Just a few of the most common ones are reviewed here.

Features

- **Modulation Measurements**
  - **time domain**
  - **FFT**
  - **AM/FM detector**
  - **time-gating**
- **Stimulus Response Measurements**
  - **tracking generator**
- **Basic Operation**
  - **remote operation**
  - **markers**
  - **limit lines**
- **Noise Measurements**
  - **noise marker**
  - **averaging**



12 – Spectrum Analysis
H7215A#101 v2.0

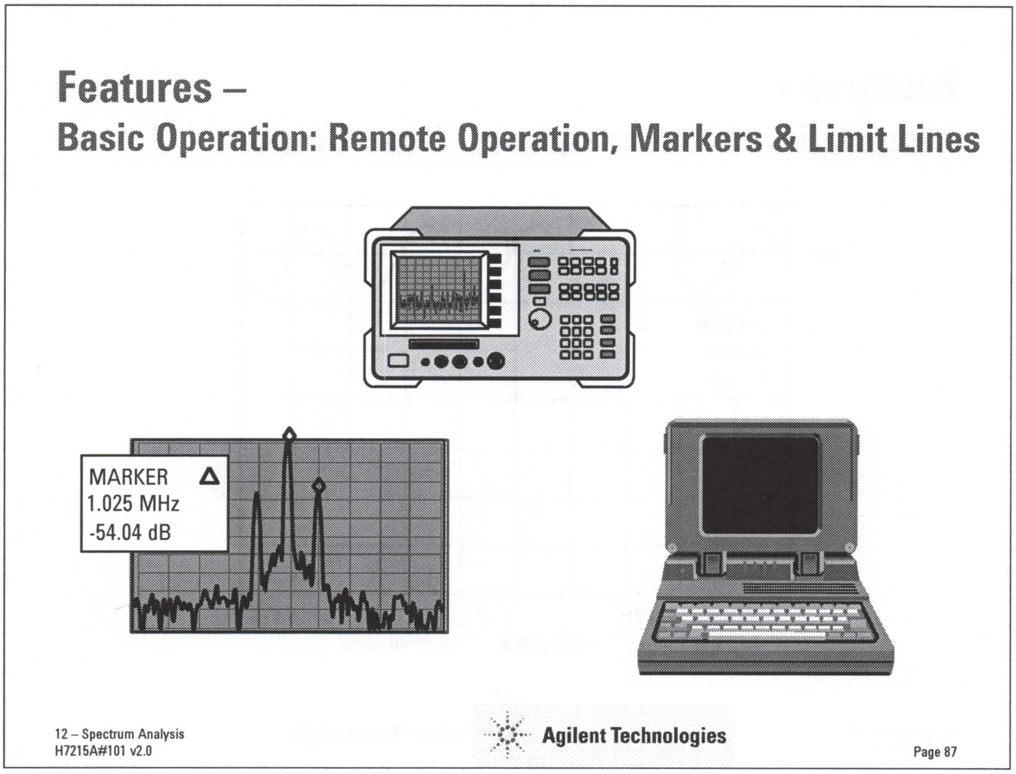
Agilent Technologies

Page 86

The features are categorized into application areas in order to better describe their function.

The first group, under Basic Operation, are some of the key features that enhance the use of the analyzer for any application.

The others refer to a specific application, although the feature is not necessarily used only in that application.



Automatic Test

For automated and remote operation, a computer(controller) is used to directly control the operation of the spectrum analyzer over GPIB*. In addition, spectrum analyzers with parallel or RS-232 capability can directly control a plotter or printer for getting hard copies of the CRT display. Analyzers with GPIB only can easily be interfaced to a parallel printer with an GPIB to parallel converter.

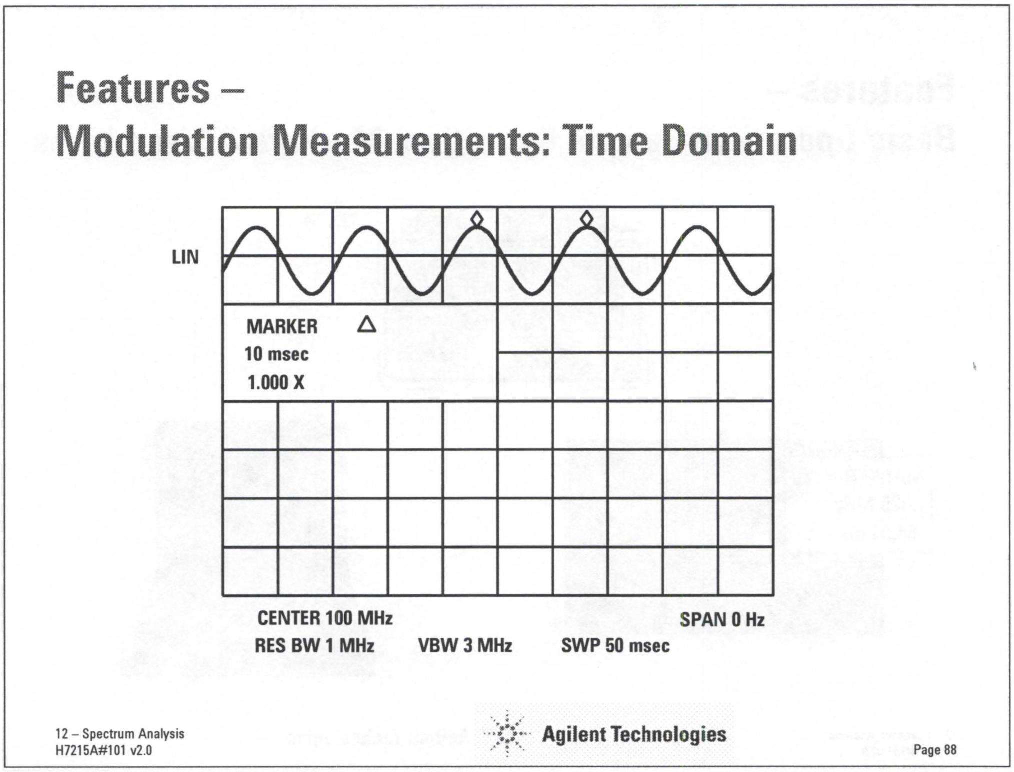
GPIB allows up to fifteen devices to be in parallel on the same bus, and be controlled by a controller which may be a computer equipped with a suitable interface or an instrument with control capability.

Markers

Markers allow quick and accurate amplitude and frequency measurements of signals, and the determination of the differences between signals. Modern spectrum analyzers provide electronic limit-line capability. This allows comparison of trace data to a set of amplitude and frequency (or time) parameters while the spectrum analyzer is sweeping the measurement range.

When the signal of interest falls within the limit line boundaries, the analyzer displays a PASS message. If the signal should fall out of the limit line boundaries, a FAIL message will appear.

* GPIB is the IEEE488 standard sometimes called GPIB.



Although a spectrum analyzer is primarily used to view signals in the frequency domain, it is also possible to use the spectrum analyzer to look at the time domain. This is done with a feature called zero-span where essentially the spectrum analyzer becomes a fixed tuned AM receiver.

Any information within the RBW is envelope detected (AM detection), the instantaneous value of the detected signal will be plotted on the display as a function of time.

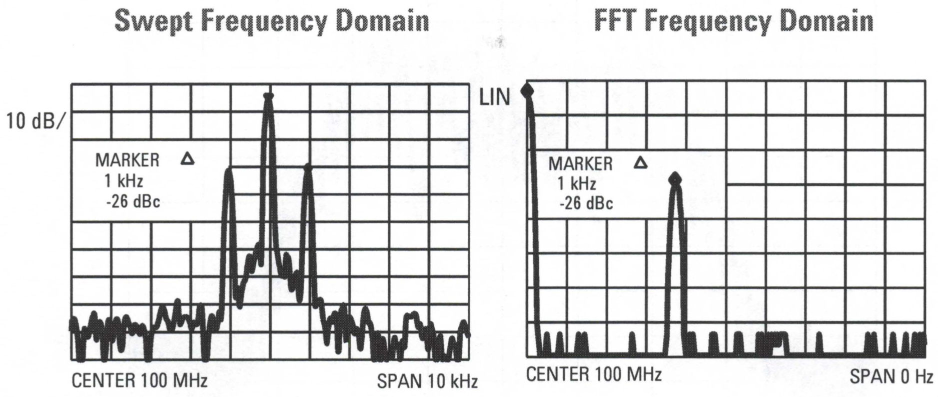
The slide is an amplitude modulated (AM) signal using zero-span.

The display shows a delta marker of 10 ms. Since this is the time between the two peaks, the period T is 10 ms. Recall that period T = 1/fmod (where fmod = modulation frequency). Hence, fmod is 100 Hz.

This feature cannot be used for quickly varying signals, since the minimum sweeptime of most analyzers is typically slower than would be necessary.

Zero-span operation is not limited to modulation measurements. It can be used to characterize any signal that is slowly varying in amplitude, such as a broadcast signal experiencing atmospheric fading.

Features –
Modulation Measurements: FFT



12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 89

The ability to do an Fast Fourier Transform (FFT) on a time-domain signal is another useful feature.

.

The graph on the left is a typical swept frequency-domain plot of an AM signal. The RBW is << modulation frequency, so the sidebands can be resolved.

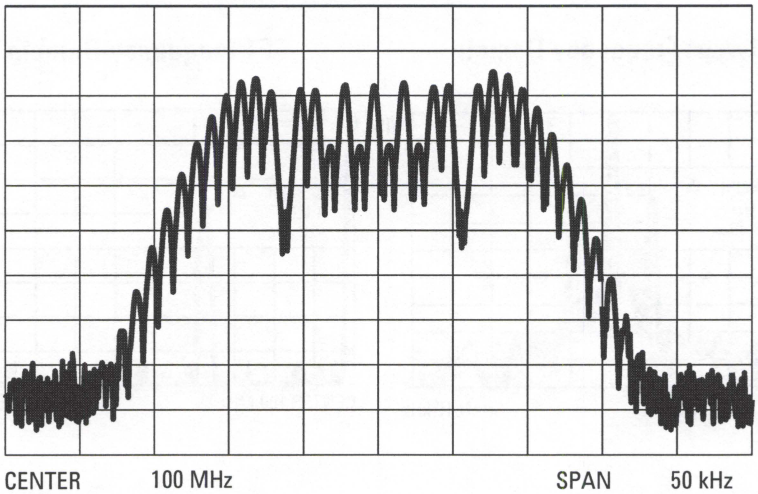
From this plot modulation frequency is determined and degree of modulation. (fmod = delta marker freq, m=2x10(A/20) where A=delta marker amplitude). For this example, fmod = 1 kHz and m = .1 or 10%

Another way to make this measurement is to use the FFT feature. While in the time domain, (now RBW>>fmod to include sidebands) . When the FFT function is selected the analyzer will usea discrete fourier transform to convert from time- to frequency-domain.

The graph on the right is an example of this feature. The carrier is at the left edge because it is at 0 Hz relative to itself.

Notice that the delta markers give us the same results as we got from the swept frequency domain.

Features – Modulation Measurements: FFT



12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 90

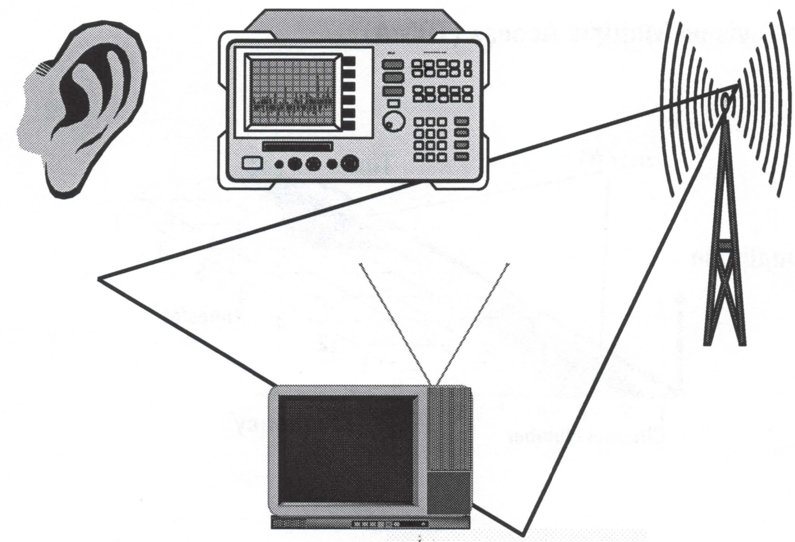
This slide shows a carrier that has both AM and FM (here again, fmod is 1kHz). The amount of FM is >> AM, so it's impossible to measure the percent AM using the swept-freq-domain. We can, however, using the FFT feature

Looking at this signal in the time-domain (zero-span) using a wide RBW, only the AM is displayed. This is because envelope detection is not sensitive to frequency deviations, only amplitude changes are detected as changes in level. Using the FFT function on this signal will give the display exactly like the one on the right of the previous slide. Now, measureing the modulation is easy.

Other advantages include better amplitude accuracy, frequency resolution and orders-of-magnitude improvement in speed. A useful measurement may be to measure line related sidebands on a microvave carrier. A swept frequency measurement with a resolution of 10Hz would be needed whereas with a resolution of 1kHz, which is the limit on many general purpose spectrum analyzers, may be used with the FFT function.

The only disadvantage is that relative frequency accuracy is not as good.

Features – Modulation Measurements: AM/FM Detector with Speakers



12 – Spectrum Analysis
H7215A#101 v2.0

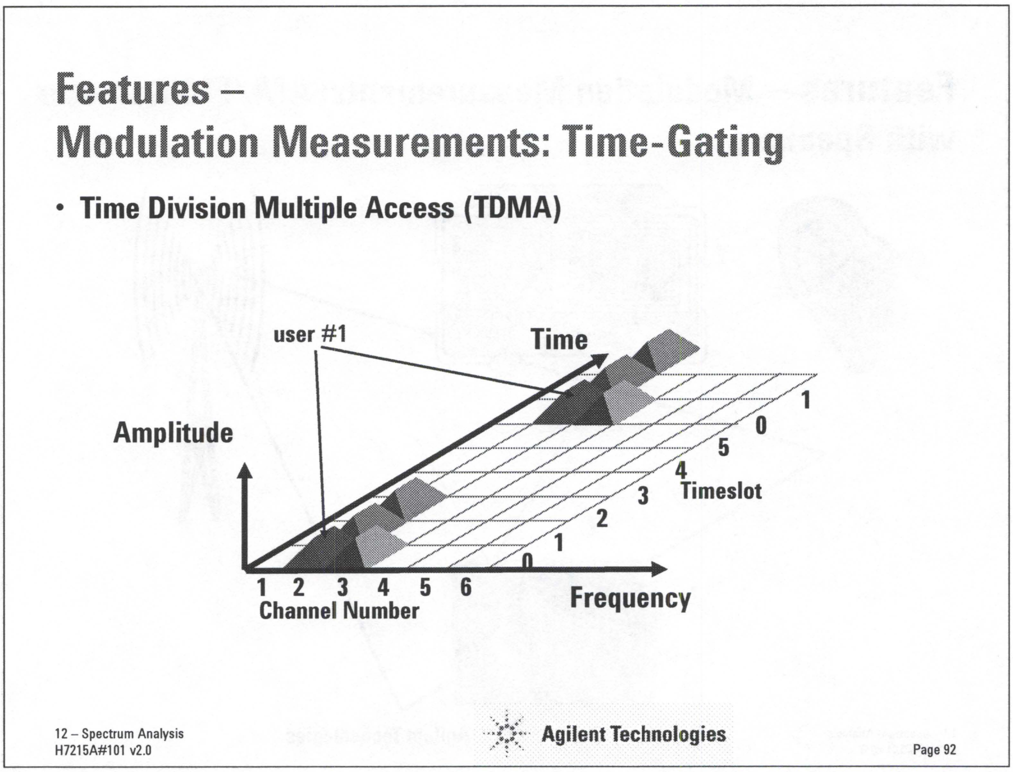
Agilent Technologies

Page 91

Most modern spectrum analyzers have AM/FM detection with audio output to a loudspeaker.

The built-in AM/FM detector with the speaker allow the modulation to be heard. An intelligible signal may be easily identified, such as AM radio station, FM radio station, TV station, amateur radio operator, etc.

Unintelligible tones and buzzes from video and control signals are also easily identified by many experienced operators.



Time Gating

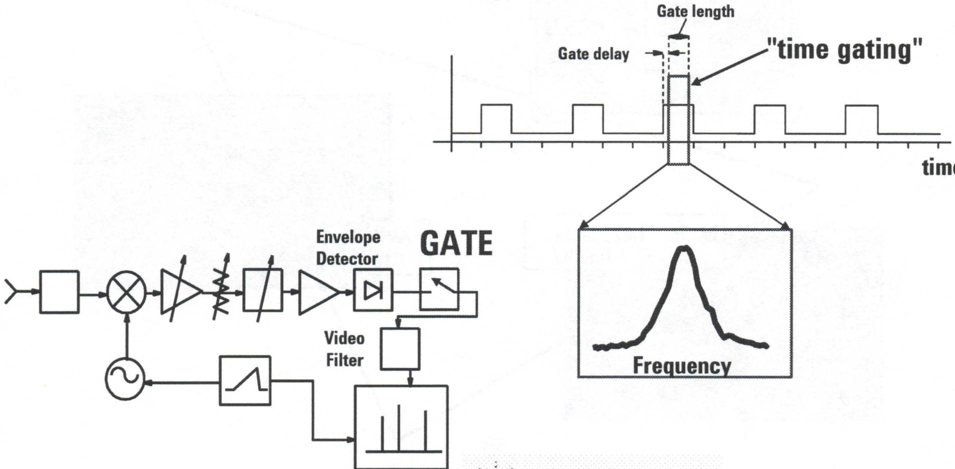
Time-Division-Multiple-Access (TDMA) is a method used to separate channels in time rather than frequency in communications systems. This is an excellent example to explain the Time Gating mode of operation. TDMA divides up the frequency channels into time slots, so that users can occupy the same frequency, but use different time slots.

In order to maintain the quality of service, verification of performance in both the time and frequency domains is necessary. The timing of the bursts, as well as the rise and fall times must be tested to verify that bursts in adjacent timeslots don't overlap.

The time-gating feature on a spectrum analyzer allows the separation of certain time variable effects, such as the change in spectrum during its switch on compared to the steady, mid-pulse spectrum.

Features – Modulation Measurements: Time-Gating

- **Time-Gated Measurements in the Frequency Domain**



12 – Spectrum Analysis
H7215A#101 v2.0

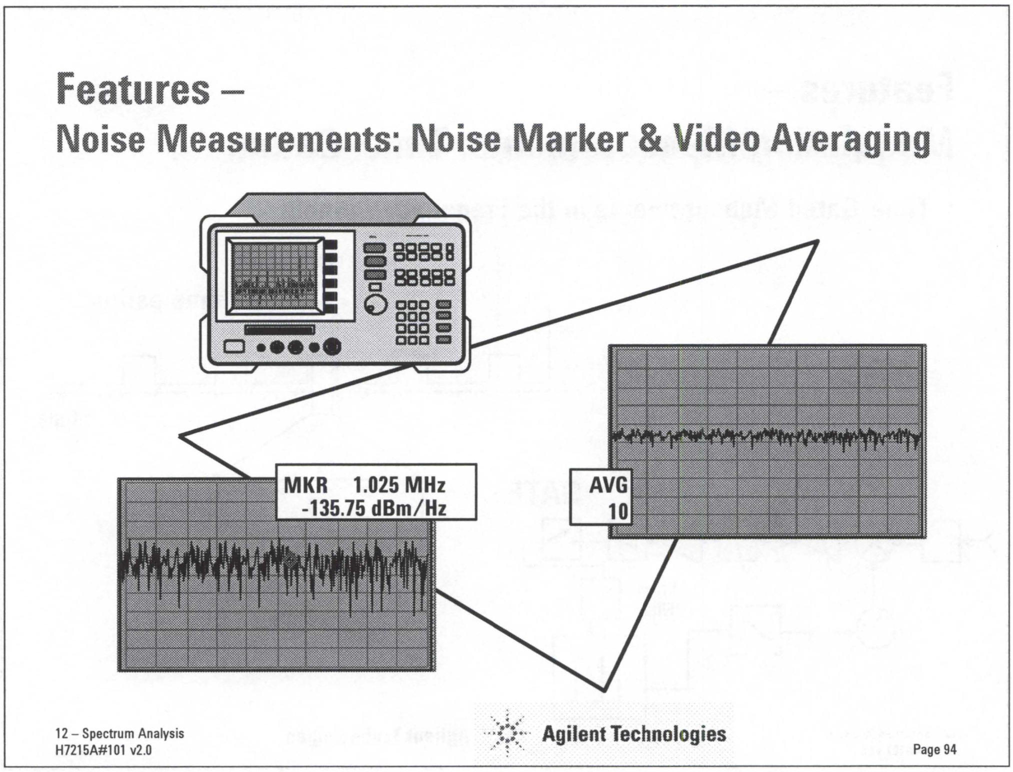
Agilent Technologies

Page 93

Time gating allows you to control the time at which the spectrum is analyzed.

The implementation in the analyzer is fairly straight forward. A video gate, or switch, is inserted between the envelope detector and the video filter. By controlling both the start of the measurement sweep (gate delay) and the duration of the measurement (gate length), the signal is allowed to reach the sampling hardware only during the selected time interval.

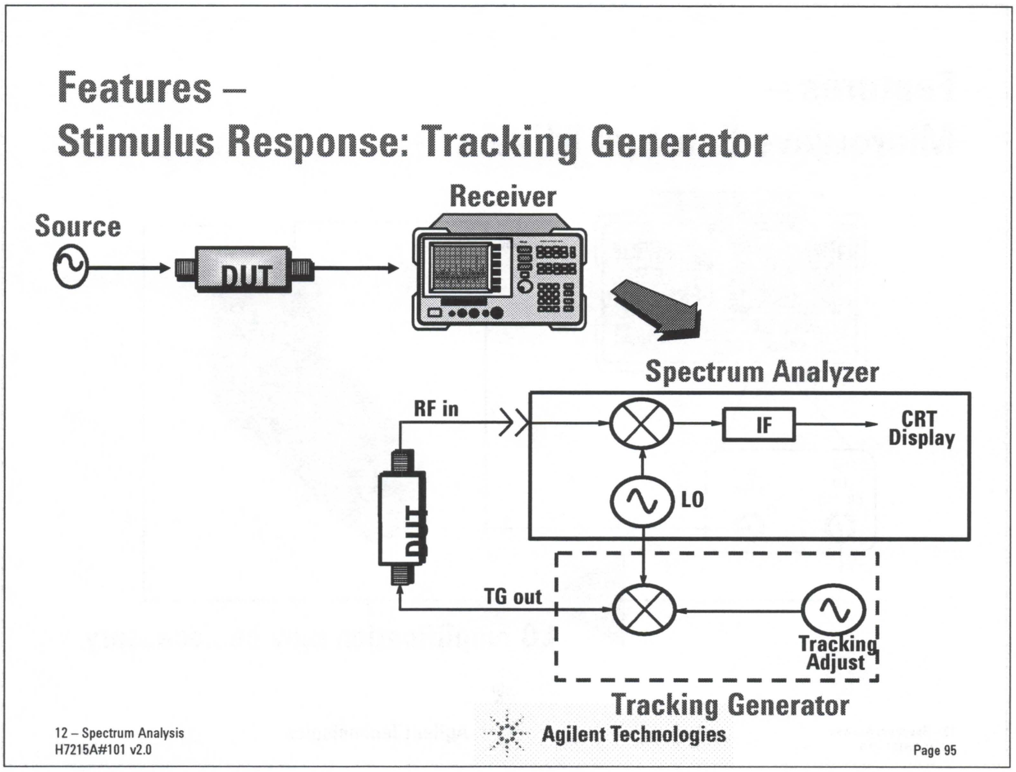
This enables the analyses the frequency spectrum during the time which the transient spectra (turning on and off of the transmitter) are present, or these may be excluded or time-filtered out so you can analyze the spectrum due to the modulation only.



When making measurements on noise, there are a some functions of a spectrum analyzer that can make the measurements easier and more accurate.

The first is a noise marker. By choosing Noise Marker as opposed to a normal marker, the value displayed as a power spectral density value in a 1-Hz noise power bandwidth. When the noise marker is selected, sample detection mode is used, as this does not impart bias to the sampled data, also several trace elements each side of the marker position are averaged, this value is corrected by a factor to account for detection, bandwidth, and log amp effects, and this value is then normalized to the 1 Hz bw. This direct reading marker is a great convenience when making noise measurements.

Another feature that is useful when making measurements on random noise, is video averaging. This is a digital averaging of a spectrum analyzer's trace information and is only available on analyzers with digital displays. The averaging is done at each point of the display independently and is completed over the number of sweeps selected. Sample display detection is automatically selected when digital averaging mode is selected.

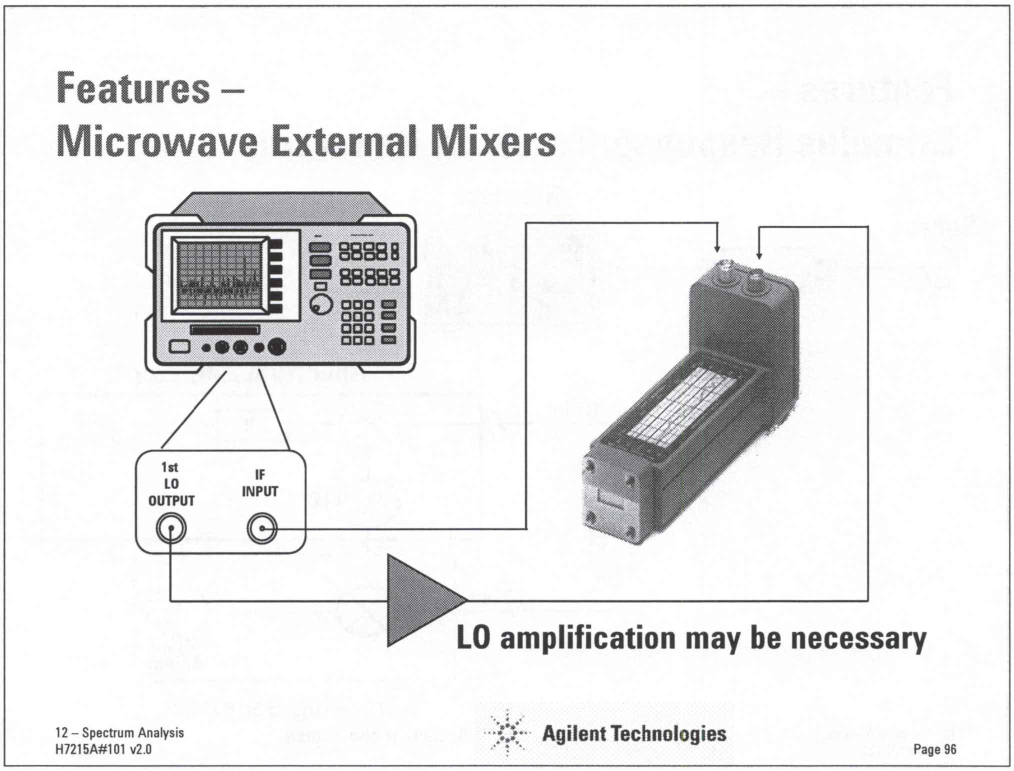


Tracking Generator

Frequency response, return loss, conversion loss, may be measured, the receiver being the spectrum analyzer, and a special source called a tracking generator is used. A tracking generator is typically built into the spectrum analyzer and is a sinusoidal output whose frequency is the same as the the tuning of the spectrum analyzer.

The output of the tracking generator (source) is connected to the input of the DUT and the response is measured by the analyzer (receiver). As the analyzer sweeps, the tracking generator will always be operating at the same frequency so the transfer characteristics of the device is measured.

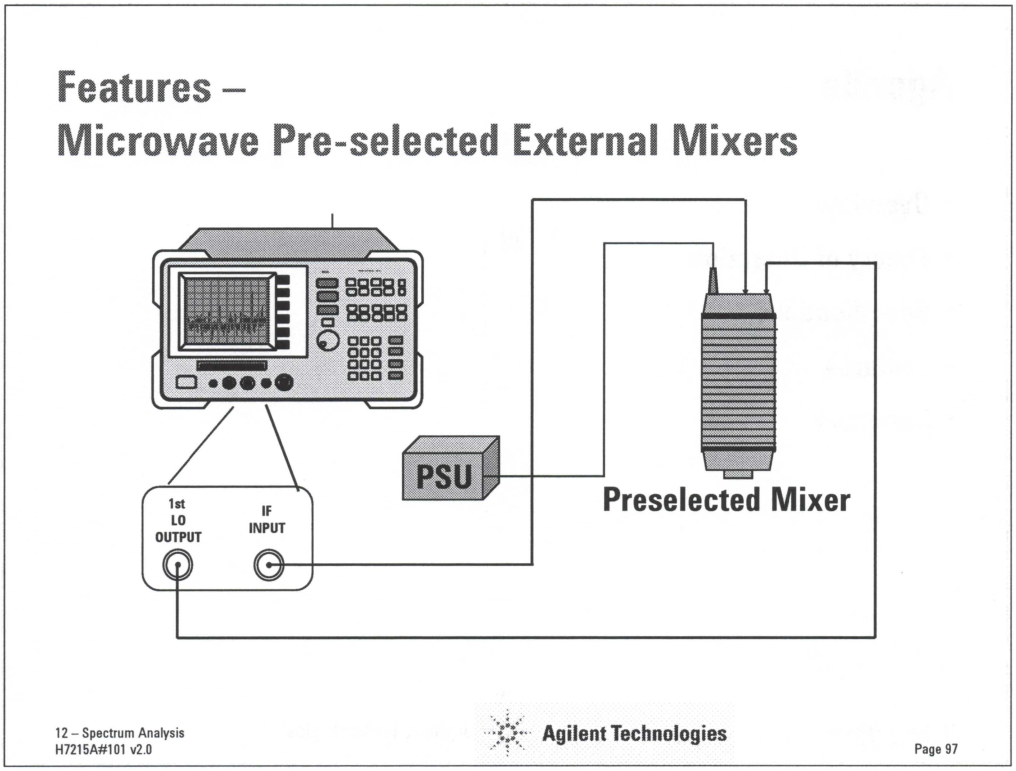
Because the spectrum analyzer can measure over a very large dynamic range, a good application for the system is screening (shielding) effectivness or isolation.



The external mixer is an accessory which allows a spectrum analyzer to measure at waveguide bands. The analyzer must have access to the LO output and an input which allows the internal mixer to be bypassed. Some older analyzers may require the LO output to be amplified (A mixer needs a strong LO signal)

Signal ID

Using a simple external mixer needs skill in signal identification.Several different and unwanted mixing responses will coincide with the IF center as a sweep is made, these will appear on the display together with the correct responses. Modern analyzers have built in signal identification tools, but with dozens of signals on the display, the identification process can be a tedious but necessary process.



The preselected mm wave mixer enables the user to make measurements quickly, without the tedium and confusion of a display that not only shows the result of mixing with calibrated harmonics but also all the others. Just as in the microwave spectrum analyzer the preselection process ensures only those mixing products which are wanted are displayed.

Agenda

- **Overview**
- **Theory of Operation**
- **Specifications**
- **Features**
- **Summary**

12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 98

The key message we hope to leave you with, is that spectrum analyzers are extremely useful tools for characterizing and analyzing a variety of devices and systems.

All it takes is a basic understanding of how they work and their characteristics, in order to use them effectively both for making accurate measurements as well as properly interpreting and analyzing results.

Books and other sources:

- Spectrum and Network Measurements.
- Robert A . Witte: Prentice Hall: ISBN 0-13-030800-5

- Application Note 150:Specrum Analyzer Basics; Agilent.
- Application Note 150-1:Amplitude and Frequency Modulation; Agilent.
- Application Note 1286-1: 8 Hints for making better Spectrum Analyzer Measurements; Agilent.

What's Important?

- What settings provide the best sensitivity?
- How do you test for analyzer distortion?
- What determines dynamic range?

12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 99

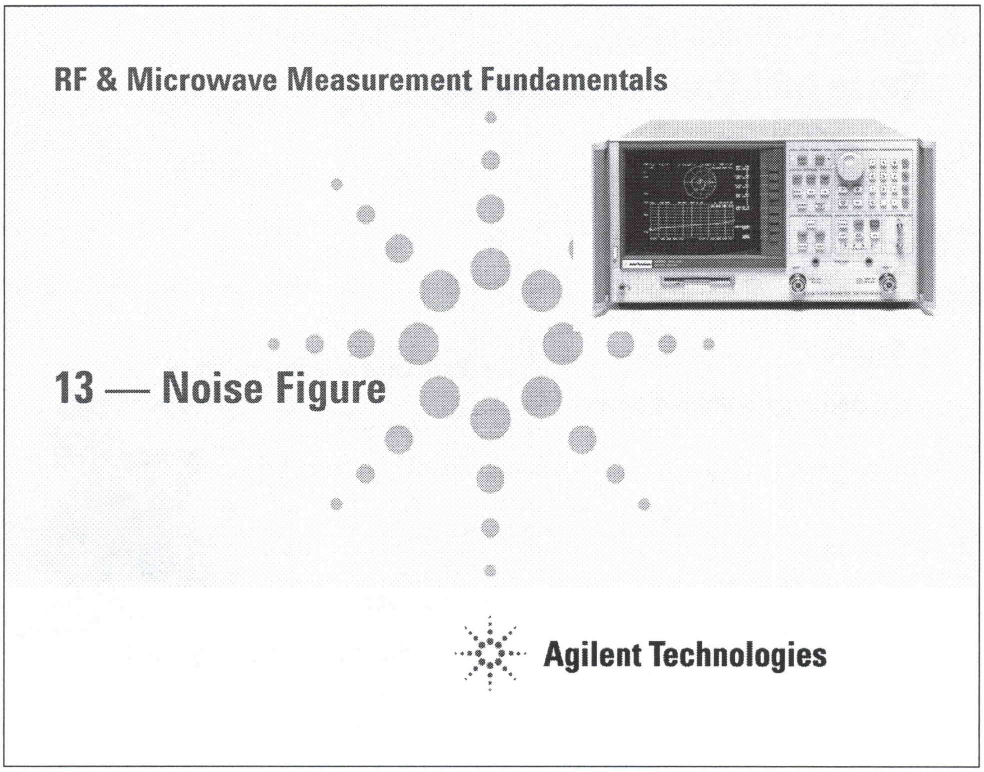
What's Important?

- **What settings provide the best sensitivity?**
  - **No input attenuation**
  - **narrowest resolution bandwidth**
  - **low level input**
- **How do you test for analyzer distortion?**
  - **Increase the input attenuation and look for signal amplitude changes**
  - **Set the attenuator at the lowest setting without change**
- **What determines dynamic range?**
  - **Analyzer distortion, noise level, and sideband noise**

12 – Spectrum Analysis
H7215A#101 v2.0

Agilent Technologies

Page 100

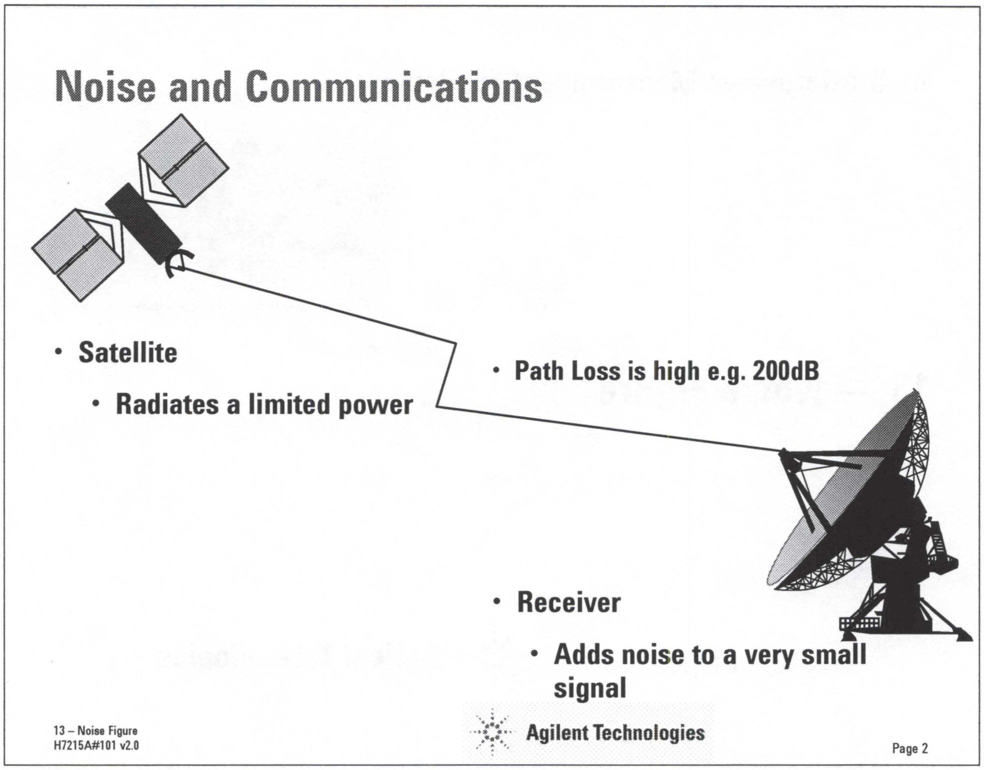


This section is to do with noise as it occurs in nature. Some disciplines use the term "noise" more loosely to include any unwanted signals that interferes with the wanted signal, e.g. crosstalk.

Noise is fundamental to our understanding of communications; this was formalized by Shannon in his seminal paper "A Mathematical Theory of Communication" BSTJ (July 1948). The famous result means that our ability to communicate in a given bandwidth is limited by noise.

$$C = W \log_2\left(\frac{P+N}{N}\right)$$

Which relates Capacity, C Bits/sec. to Bandwidth W , Mean signal power P, and mean noise power N. If noise N can be reduced then we may also reduce bandwidth for the same capacity. In the limit, N → 0, an infinite amount of information may be sent over an arbitrarily small bandwidth. Of the types of noise above we shall be dealing with thermal noise.



A satellite to ground microwave link is a good example of the importance of understanding noise processes. Satellites can radiate a limited amount of power usually in the order of 10W. The satellite antenna system is designed to project this power to the region where the receiving antenna is located. The directional property of both the transmit and receive antennae is quantified by their respective antennae gains. Nevertheless the received signal power is very small due to the high path loss. The receive antenna not only receives the transmitted signal but also interference in the form of noise. The receiver also contributes noise of its own, in the receiver channel, the signal to total noise ratio is a critical parameter. A poor S/N can spoil or even swamp the the received information, for example "snow" on a TV picture.

We shall describe the origin of noise, how it is quantified and expressed, and how to measure it's effect.

Noise Figure Topics

- Noise Processes and Definitions
- Thermal Noise and Noise Power
- Noise Figure
- Noise Temperature
- The Importance of Noise Figure
- Noise Figure Measurement

13 – Noise Figure
H7215A#101 v2.0

Agilent Technologies

Page 3

Noise Processes

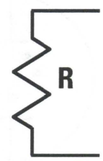
- Thermal, Johnson, Nyquist
- Shot
- Flicker ($1/f^{n}$)
  - And many others

Johnson and Nyquist are the names of scientists/engineers involved.
Nyquist's name occurs (same guy) in sampling and feedback theory. Nyquist and Shannons names are linked in sampling theory.
NOISE LIMITS OUR ABILITY TO SEND INFORMATION IN A GIVEN BW

Electrical noise from natural sources are due to different mechanisms, thermal noise will be the subject of the consideration for this paper.

Thermal Noise

- **First suggested by Shottky in 1918 as a noise generated by random thermal motion of electrons in a resistor.**
- **Measured by J.B Johnson of Bell Labs 1928.**
- **Theoretical foundation done by H. Nyquist also of Bell Labs.**



$$\overline{V^2} = 4kRTB$$

(This gets simpler later.)

13 – Noise Figure
H7215A#101 v2.0

Agilent Technologies

Page 5

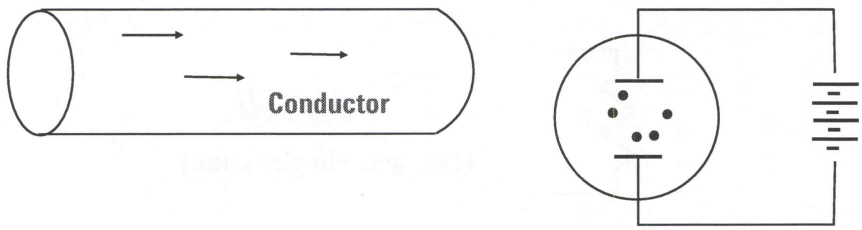
This noise is called variously Thermal, Johnson, and sometimes Nyquist noise. Nyquist developed the above expression in terms of k, Boltzmans constant, R, resistance, T, absolute temperature and B the bandwidth.

We will return to this subject later when the available power from any passive source will be developed.

The absolute temperature scale has its origin at absolute zero (– 273° C). The scale has the same size degrees as the Celsius scale and is designated with K. Thus 273K is 0° C, and 25° C is 298K.

Shot Noise

- **Motion of individual electrons are considered as independent events.**
- **Spectral density proportional to current and independent of temperature (unlike thermal noise).**



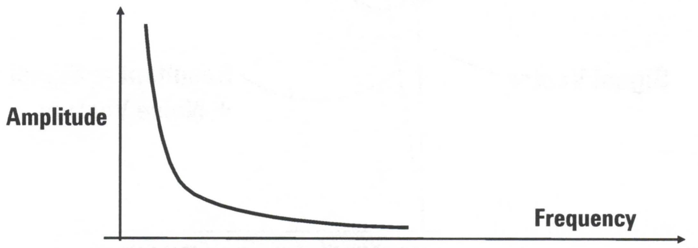
Shot noise is caused by the quantitized and random nature of current flow. The smallest unit of charge (one electron) is $e = 1.6 \times 10^{-19}$ coulombs. When a DC current I_0 flows, the average current is I_0 , but there are fluctuations in amplitude which have a spectral density which is fairly uniform with frequency.

$$I_n^2(f) = 2eI_0 A^2/Hz.$$

Thermal and shot noise are the types with witch we are concerned when measuring noise figure.

Flicker Noise

- **Sometimes called 1/f noise**
- **Becomes significant at very low frequencies or at frequencies very close to a carrier in oscillators.**

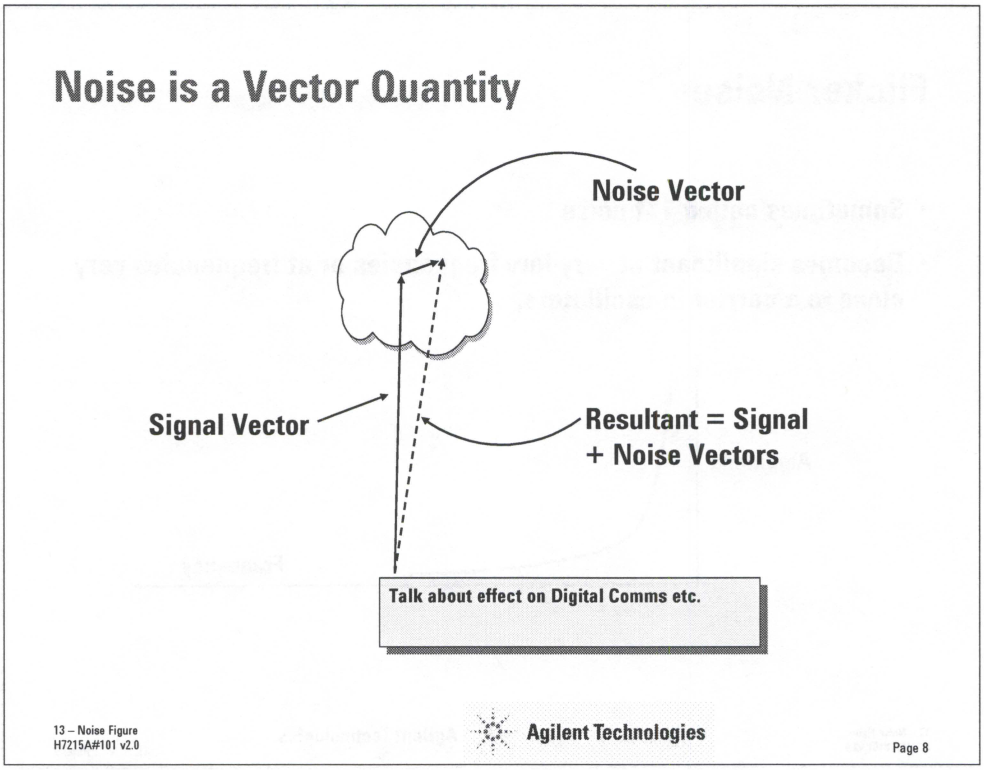


13 – Noise Figure
H7215A#101 v2.0

Agilent Technologies

Page 7

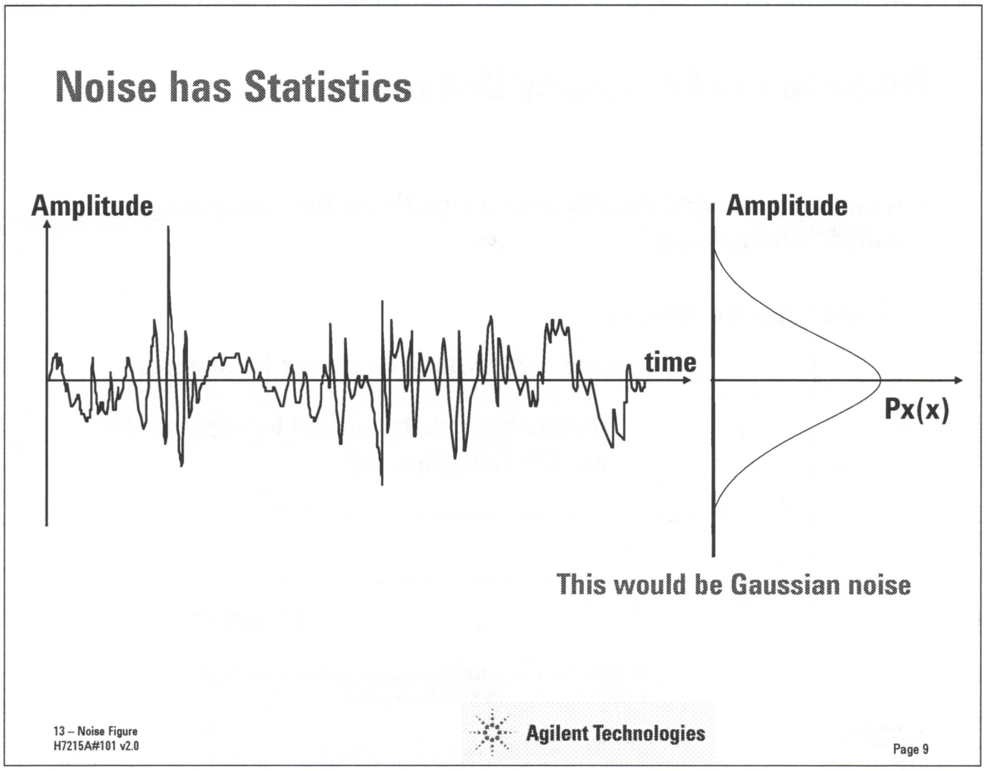
Flicker noise is any noise whose amplitude varies inversely with frequency.



When noise is added to a sine wave they add vectorially; in other, words not only are there amplitude fluctuations but also phase fluctuations in the resulting signal.

In simple communication systems, e.g. using linear amplitude modulation, noise always degrades the received information.

In a modern (digital) communications system, whether transmitted symbols are coded by phase or amplitude or both, added noise will increase the probability of error in the detection of those symbols. However, no matter what the signal-to-noise ratio is , as long as the symbols are interpreted properly by the receiver, the original signal can be reconstructed without degradation.



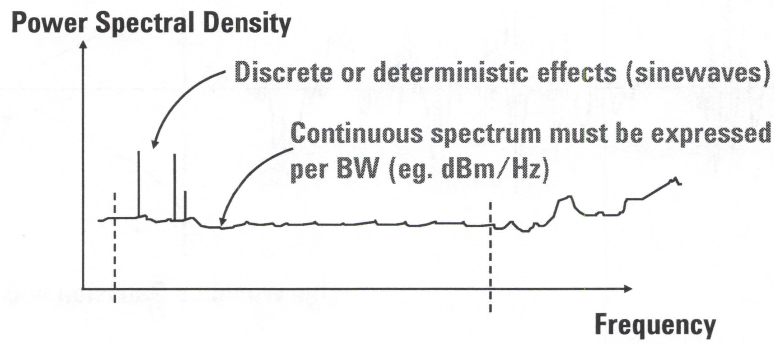
A noise voltage is random; one cannot predict its value at any time from it's previous behavior. Noise, however, may be described by statistical parameters; if, for example, the PDA (Probability Density of Amplitude) is approximately a normal distribution we may call the noise Gaussian. This is usually the case.

Whatever the PDA, the classic statistical parameters, Standard Deviation and Variance, have some physical meaning. If the unit of the measured noise amplitude is the volt, then the SD would be the root-mean square value ;*Vrms*, and the Variance the power(watts) in a 1 ohm resistor.

Another important statistical parameter is *correlation*; this is the measure of similarity, in this case between two noise sources. Independent noise sources are usually uncorrelated. The practical importance of this relationship is that noise sources, when combined, add in *power.* Perfectly correlated sources would add in *voltage.*

Noise in the Frequency Domain

- **A constant spectral density over a significant frequency range is called "White Noise."**



White noise is so called because it's analagous to White Light, its got all the frequencies

13 – Noise Figure
H7215A#101 v2.0

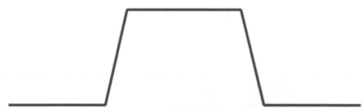
Page 10

A term which is associated with the frequency domain view of noise is power spectral density; this is an important concept. Since noise may be thought of as being made up of an infinite number of spectral lines arbitrarily close together. Then power must be measured in a bandwidth. Such units as dBm/Hz are used. If the spectral density is reasonably flat over a given band then the noise is white noise, pink noise has constant energy per octave.We shall assume white noise in this paper.

It is important to distinguish between noise which must be expressed as a power density and deterministic signals (sine waves). The measurement of a sine wave does not depend on bandwidth.

Spectral Density

- **The bandwidth associated with noise power spectral density is NOT the 3dB bandwidth.**



13 – Noise Figure
H7215A#101 v2.0

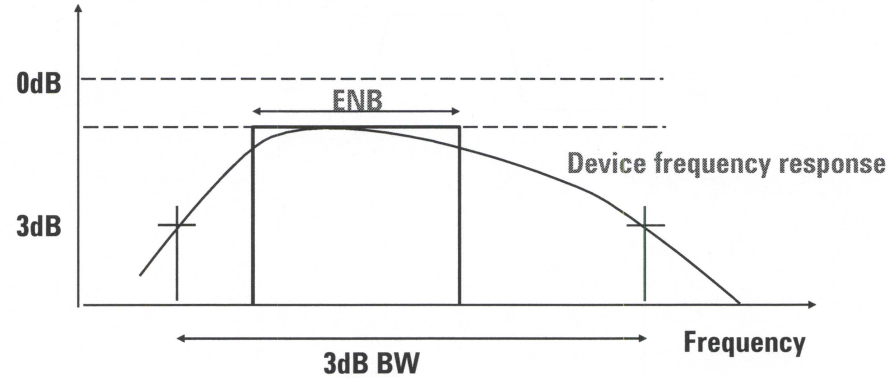
Agilent Technologies

Page 11

It cannot be overemphasized that spectral density in units of power/BW is very important. Noise power without a statement of the associated bandwidth is meaningless. This bandwidth is called the *noise bandwidth* or *equivalent noise bandwidth, or noise power bandwidth.* This bandwidth is not the usual 3dB bandwidth of filters such as those found on spectrum analyzers.

Bandwidth is Equivalent Noise Bandwidth

- **ENB is the BW of a perfect bandpass device that would pass the same noise power as the real device.**



13 – Noise Figure
H7215A#101 v2.0

Agilent Technologies

Page 12

ENB or just Noise Bandwidth is an equivalent rectangular pass band that passes the same noise power as the actual system being considered. For most filters noise BW is close to the 3dB bandwidth. In Agilent spectrum analyzers the noise BW is about 1.05 to 1.15 times the 3dB. Bandwidth.

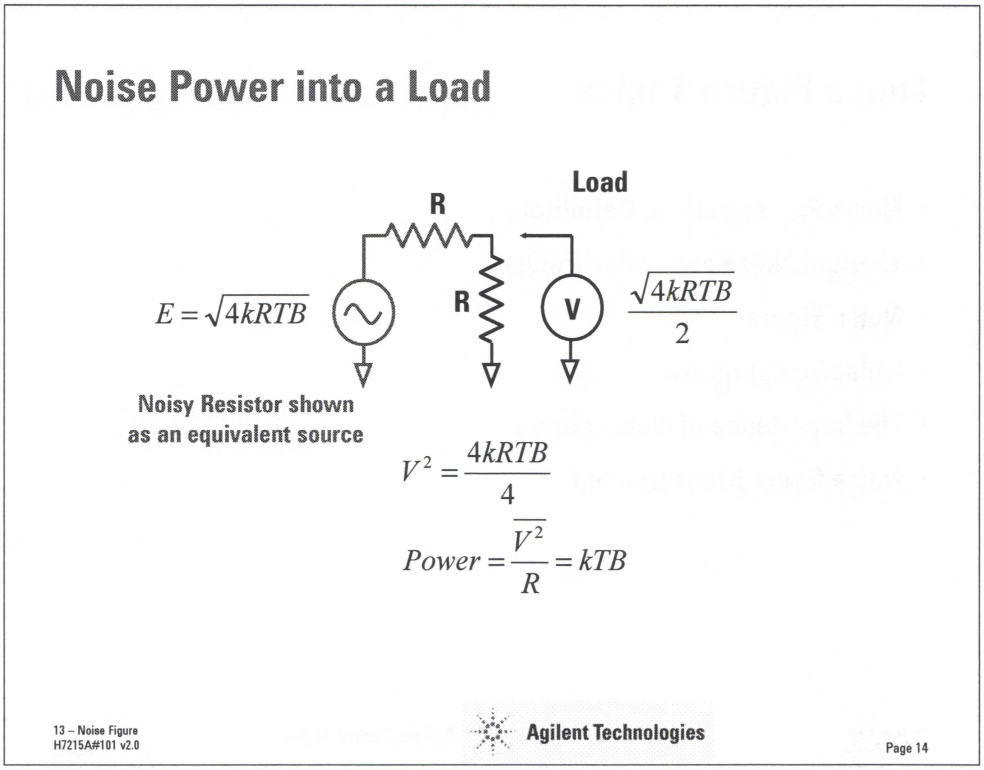
Noise Figure Topics

- Noise Processes and Definitions
- Thermal Noise and Noise Power
- Noise Figure
- Noise Temperature
- The importance of Noise Figure
- Noise figure Measurement

13 – Noise Figure
H7215A#101 v2.0

Agilent Technologies

Page 13

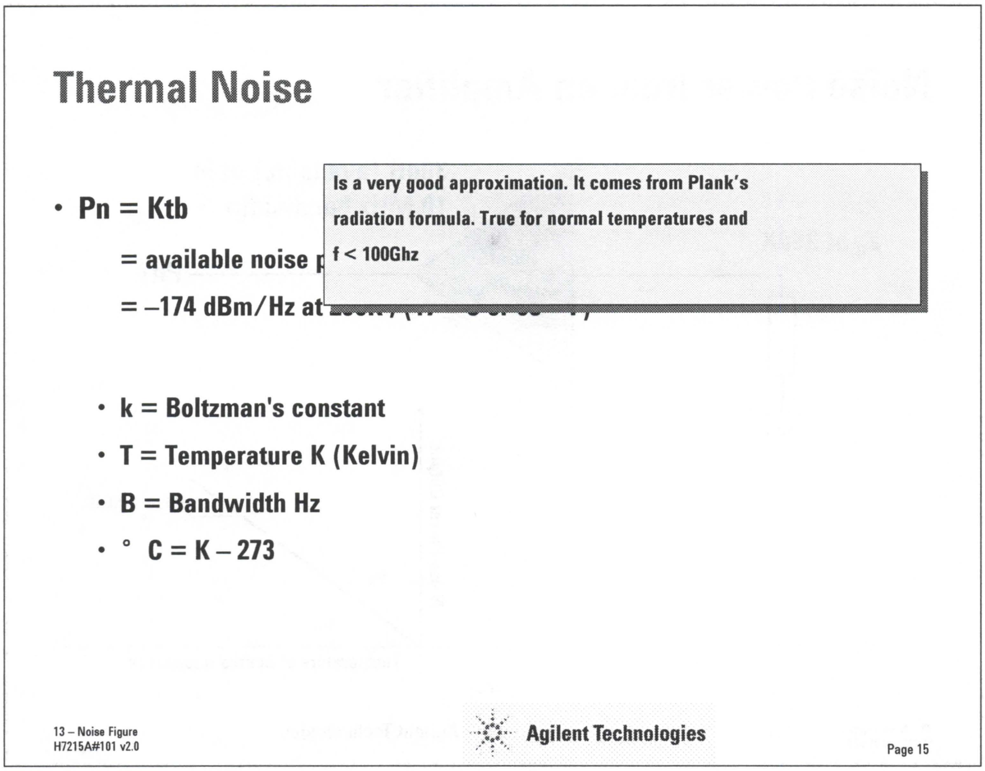


This expression of noise voltage in terms of resistance, absolute temperature, bandwidth and Boltzmans constant was stated earlier; now this is reduced to:

$$P = kTB.$$

Power (Watts) = Boltzmans Constant × Absolute Temperature (K) × Bandwidth (Hz)

In the above figure note that the load resistance is made equal to the resistance of the source resistance. This simple noise power formula is based upon complex conjugate power transfer.



This is a very significant formula. Notice there is no resistance or impedance term, this is because the term *available* power implies conjugate matching between the source and load. When dealing with noise, be aware of the way in which power transfer is defined.

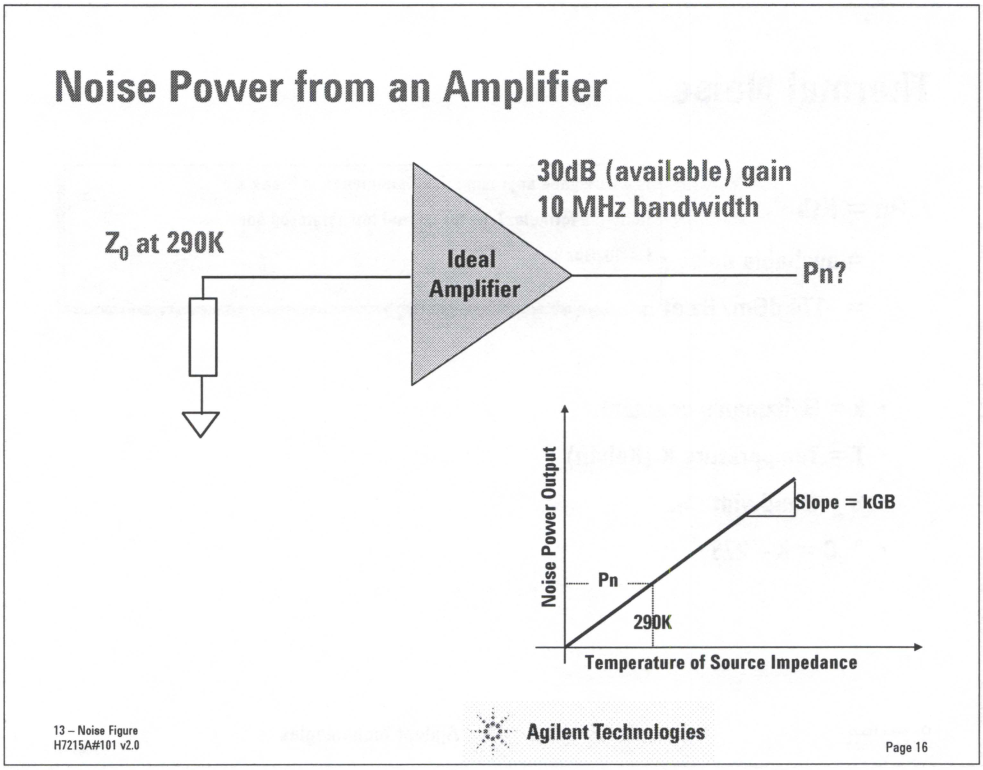
T_0 = 290K, is the *standard temperature*, this is accepted as the noise temperature seen by a receive antenna for terrestrial communications. This is also called "room temperature", which at 17° C or less than 63° F, is a cool room.

$$\text{Boltzman's Constant} = 1.38 \times 10^{-23}\ \text{J/K}$$

$$290\ \text{K} \times 1.38 \times 10^{-23}\ \text{J/K} \times 1\text{Hz} = 4.002 \times 10^{-21}\ \text{J/sec (Watts)}$$

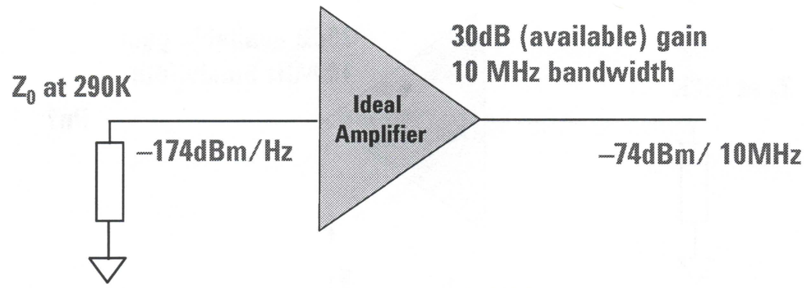
multiply by 1000 to convert to mW and calculate the dBm.

$$10 \times \log(4.002 \times 10^{-21} \times 1000) = -173.977\ \text{dBm}$$



If a broadband power meter were connected to the output, what would it read? The arithmetic is simple and does not need a calculator.

Noise Power Calculation



Pn = (PkT x 1000 x 10 000 000) mW

= Power in x gain x bandwidth ratio

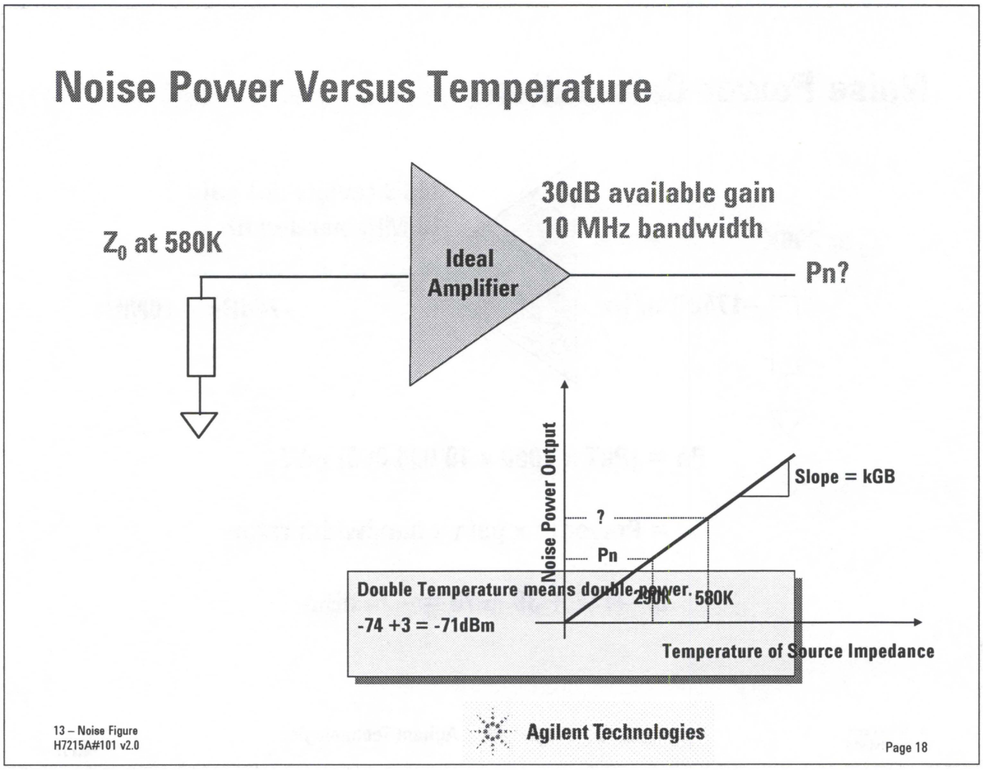
or **– 174 + 30 + 70 = – 74 dBm**

13 – Noise Figure
H7215A#101 v2.0

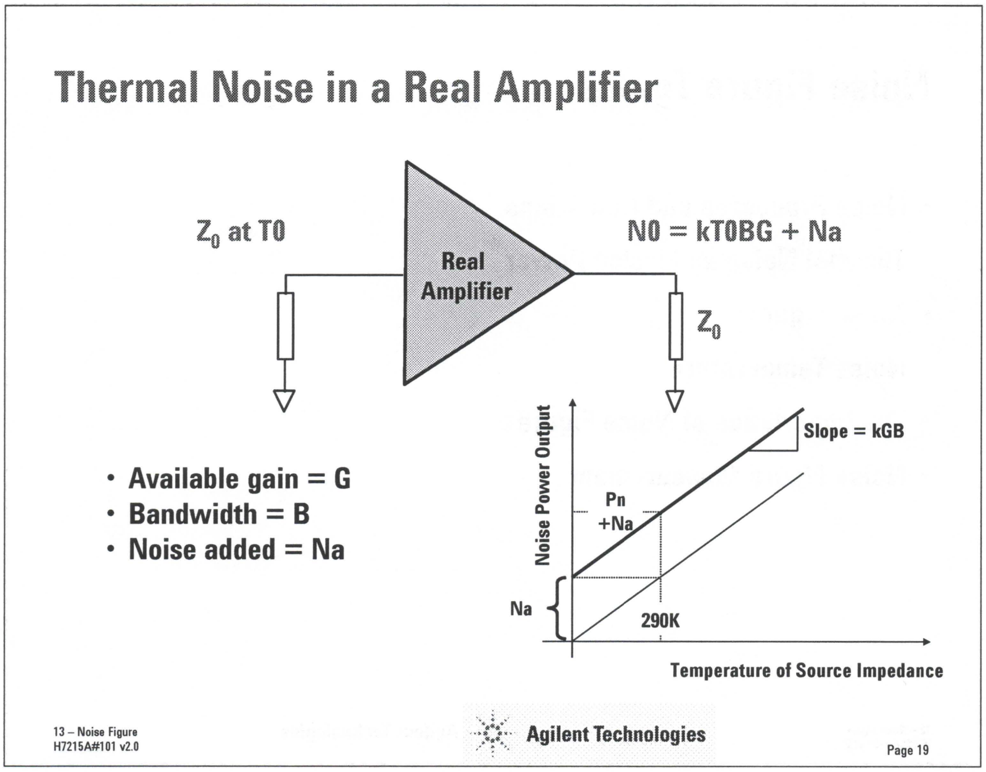
Agilent Technologies

Page 17

Input spectral power density =	–174dBm/Hz
This will be greater by ×10,000,000 or	+70dB [10log(10MHz/1Hz)]
Also by the gain	+30dB
	–74dBm/10MHz



What happens to the output power if the source termination is doubled in temperature?



In our previous examples we assumed a perfect amplifier, it added no noise to the process. Here we show that at the output there is not only the amplified input noise but also some added noise due to the amplifier components and processes.

Noise Figure Topics

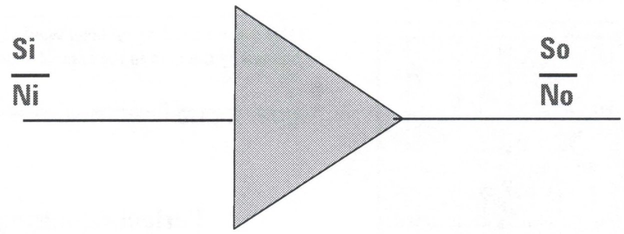
- Noise Processes and Definitions
- Thermal Noise and Noise Power
- Noise Figure
- Noise Temperature
- The Importance of Noise Figure
- Noise Figure Measurement

13 – Noise Figure
H7215A#101 v2.0

Agilent Technologies

Page 20

Noise Figure



$$\frac{Si}{Ni} , \frac{So}{No} = \text{Noise figure}$$

$$\frac{Si}{Ni} = \frac{So}{No} \quad \text{Ideal case: noise figure = 1}$$

13 – Noise Figure
H7215A#101 v2.0

Agilent Technologies

Page 21

Any noise added by the amplifier degrades the signal to noise ratio from input to output. The ratio of the s/n ratios is the noise factor or figure*.

$$F = \left.\frac{(S_i / N_i)}{(S_o / N_o)}\right|_{T_s = T_0} = \frac{S_i / N_i}{GS_i / N_o} = \frac{1 / N_i}{G / (GN_i + N_a)}$$

$$= \frac{N_a + GN_i}{GN_i}$$

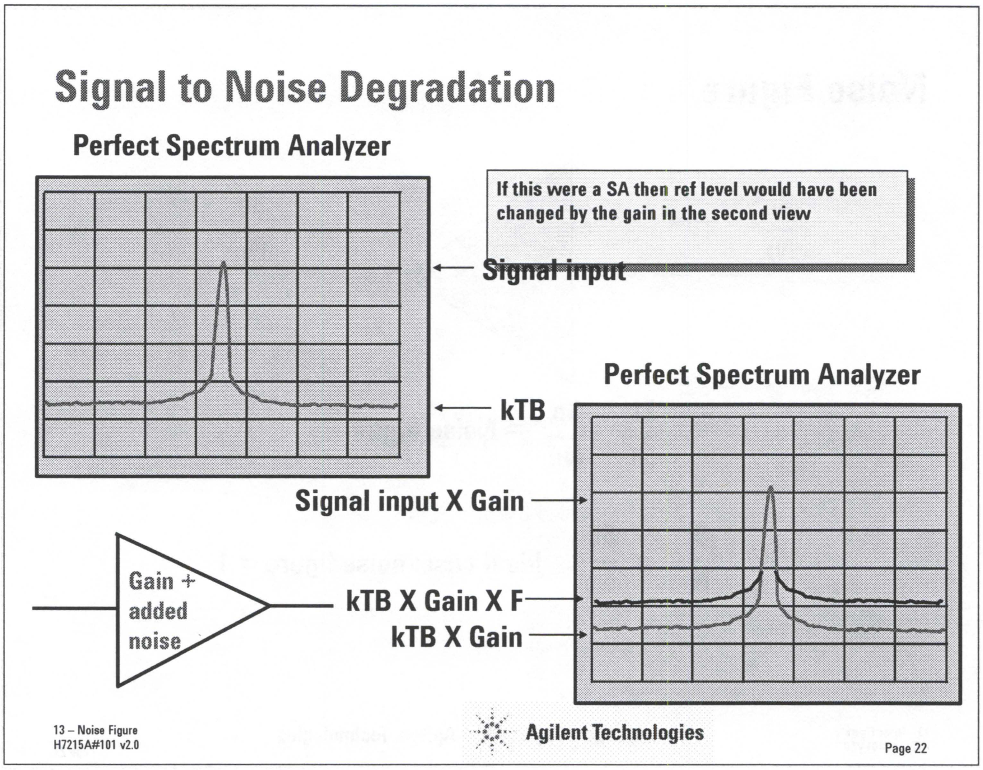
The noise at the input $N_i = kTB$, so

$$F = \frac{N_a + kT_0BG}{kT_0BG}$$

10.log(F) = 10.log(Na + kT0BG) – 10.log(kT0BG)

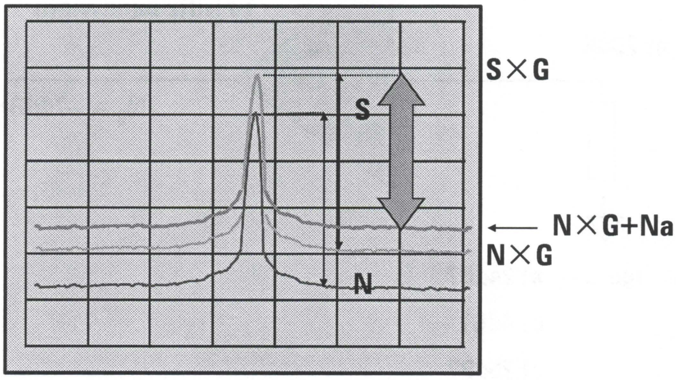
FdB = Actual output noise(dBm) – Noise from "perfect" device(dBm).

*In general the term noise figure and noise factor are synonymous, however the term noise figure is sometimes used for the ratio expressed in dB.



Here is another picture that may help the understanding of Noise Figure as a factor involved with signal to noise degradation.

Signal to Noise Degradation

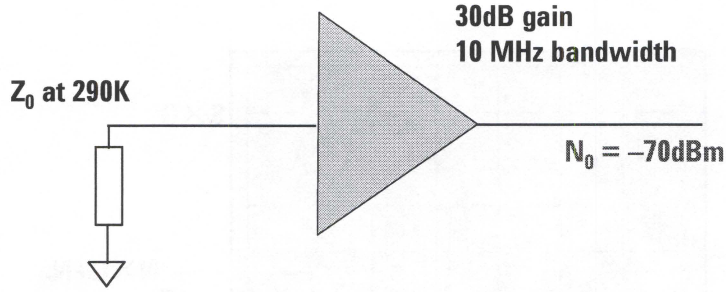


13 – Noise Figure
H7215A#101 v2.0

Agilent Technologies

Page 23

What is the Noise Figure?



- Noise Figure = a) 24dB?
  b) 4dB?
  c) 25dB?
  d) 6dB?

b) this is based on previous example NF=0dB; P_0= -74dBm

13 – Noise Figure
H7215A#101 v2.0

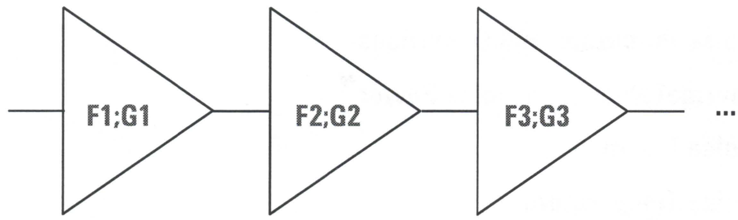
Agilent Technologies

Page 24

Which is the noise figure?

Remember: F_{dB} = Actual output noise(dBm) – Noise from "perfect" device(dBm)

Noise Figure of a Cascaded System



$$F_{sys} = F1 + \frac{F2 - 1}{G1} + \frac{F3 - 1}{G1G2} \cdots$$

The first stage dominates!

Noise figure and gain of first stage is very important:

13 – Noise Figure
H7215A#101 v2.0

Page 25

In a cascaded system of amplifiers, the niose figure of the group is dominated by the first stage, especially if that stage has significant gain.

The calculation must be done in linear units, and then converted to dB's if necessary.

Example

A spectrum analyzer has a noise figure of about 26dB. A preamp gain 26 dB and noise figure 7dB is used. What is the system noise figure?

- G_1 = 26dB. Ratio = 500
- F_2 = 26dB. Ratio 500
- F_1 = 7dB. Ratio 5

- F_{sys} = 5 + (500– 1)/500
- $\approx$ 6 as a linear ratio or 7.8dB

Noise Figure Topics

- Noise Processes and Definitions
- Thermal Noise and Noise Power
- Noise Figure
- Noise Temperature
- The Importance of Noise Figure
- Noise Figure Measurement

13 – Noise Figure
H7215A#101 v2.0

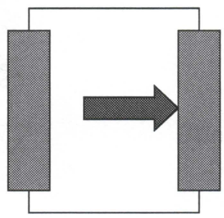
Agilent Technologies

Page 26

The idea and use of the term noise temperature is often found in papers dealing with the behavior of devices or the noise received by an antenna. Noise temperature is an absolute measure of noise power density whereas noise figure is a relative measure.

Noise Temperature Of a Passive Device

- **The Noise temperature of a resistor is the temperature of the resistor expressed in degrees K.**
- **The noise temperature of a diode may be many times the physical temperature of the diode.**



Power available (P) = kT/Hz
Noise Temperature = P/k

13 – Noise Figure
H7215A#101 v2.0

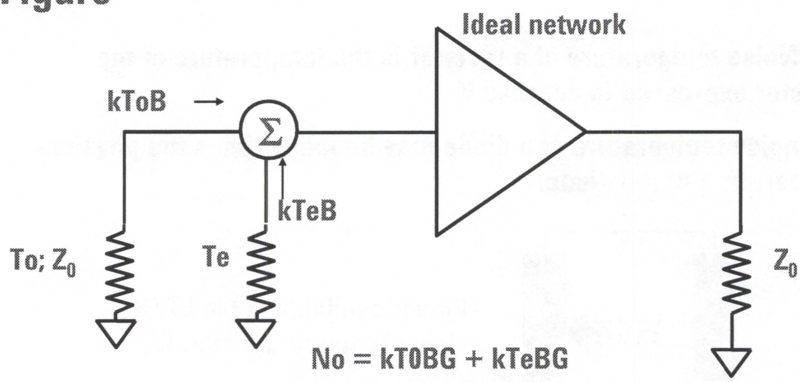
Agilent Technologies

Page 27

If Boltzman's constant is 1.38×10^{-23} W/Hz °K

And To is 290 K then kT = 4×10^{-21} W/Hz = –204dBW/Hz = –174 dBm/Hz

Noise Temperature - An Alternative to Noise Figure



$$N_o = kT_0BG + kT_eBG$$
$$= k(T_0 + T_e)BG$$

13 – Noise Figure
H7215A#101 v2.0

Agilent Technologies

Page 28

Noise temperature is another way to describe noise added by a non-ideal network. Voltage from an imaginary noise source at temperature Te is added to the noise from the actual input termination to produce the measured No.

Te is properly called the Effective Input Noise Temperature.

$$F = \frac{N_0(\text{actual})}{N_0(\text{ideal})} = \frac{k(T_0 + T_e)BG}{kT_0BG} = \frac{T_0 + T_e}{T_0}$$

so

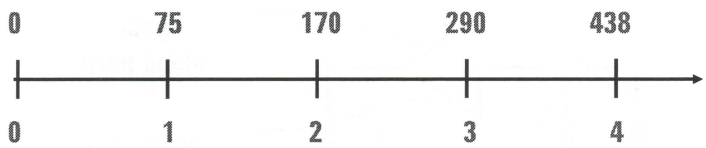
$$T_e = FT_0 - T_0 = (F-1)T_0$$

$$F_{dB} = 10\log(1 + T_e / T_0)$$

This parameter is much more useful for low noise devices such as found in satellite earth stations. Te is preferred for low noise devices because it is a more sensitive indicator when Na is small. The 290 K standard temperature is not a good reference for deep space. Te has no reference, and its use simplifies the mathematics because the noise added as well as the input noise are in terms of temperature.

Relationship between Noise Figure and Noise Temperature

- **Effective noise temperature T_e Kelvin**



Noise Figure (NF) Db

NF = 10.log(1 + Te/T0) dB

13 – Noise Figure
H7215A#101 v2.0

Agilent Technologies

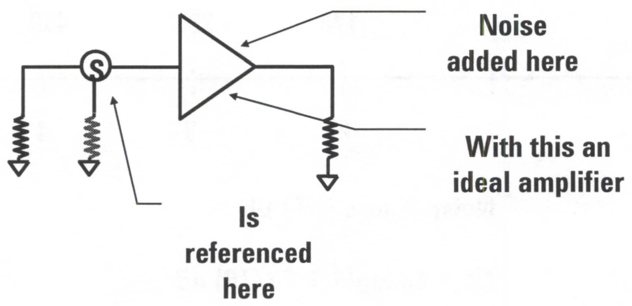
Page 29

Some more Equivalent Noise Temperatures

NFdB	Temp. K	NFdB	Temp. K
0	0	5.5	739
0.5	35	6	865
1	75	6.5	1005
1.5	120	7	1163
2	170	7.5	1341
2.5	226	8	1540
3	289	8.5	1763
3.5	359	9	2014
4	438	9.5	2295
4.5	527	10	2610
5	627		

Noise Temperature

- **Noise temperature is not a physical temperature.**



13 – Noise Figure
H7215A#101 v2.0

Agilent Technologies

Page 30

When expressing noise figure in terms of temperature referenced to the input of the device, it is important to realize that this is not a physical temperature. A 4dB noise figure is equivalent to an effective noise temperature of 438 K which is a physical temperature of 165 °C, which would feel *very hot*!

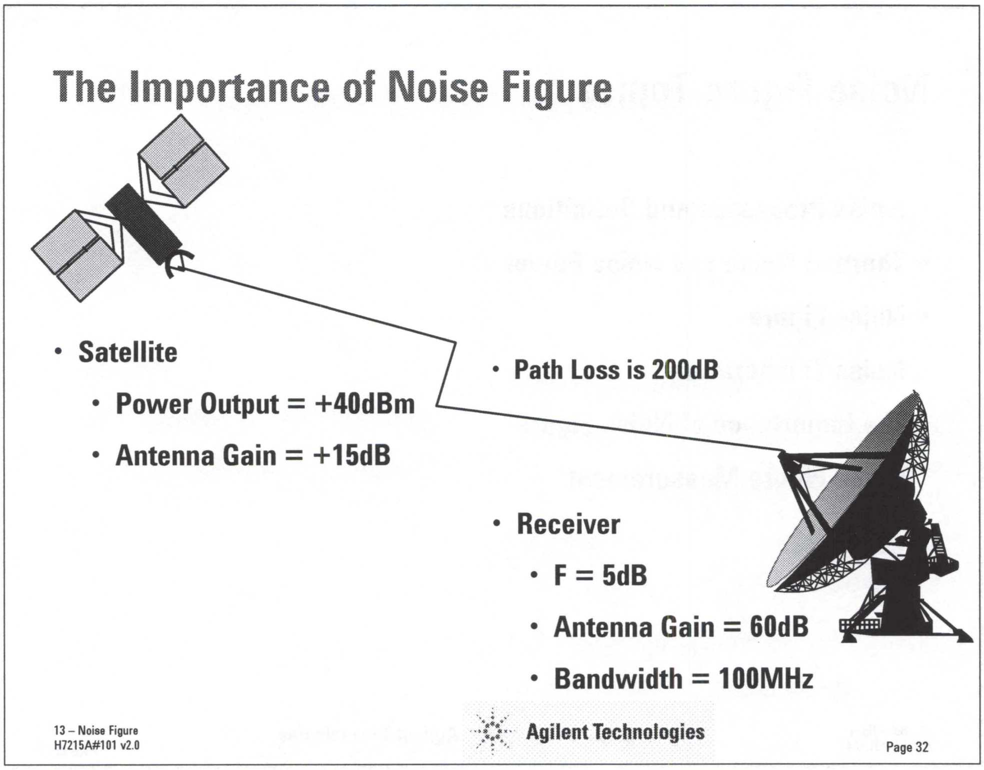
Noise Figure Topics

- Noise Processes and Definitions
- Thermal Noise and Noise Power
- Noise Figure
- Noise Temperature
- The Importance of Noise Figure
- Noise Figure Measurement

13 – Noise Figure
H7215A#101 v2.0

Agilent Technologies

Page 31



Calculations like this are done frequently in studies of satellite communication systems.

Calculate the noise margin in the above example. The receiver needs a S/N of at least **6dB** to function well, will the above system work properly? Assume that the noise temperature at the receiver (at the focus of the receive antenna) is T_0, the standard noise temperature 290K. *Noise is **not** magnified by antenna gain.*

Noise and Power Budget of Satellite Receiver - Values in dB

Signal Power to Receiver		Noise Power to Receiver	
Power Out	40dBm	kT_0	–174dBm
Antenna Gain	15dB	NF	5dB
Path Loss	–200dB	Bandwidth	80dB
Antenna Gain	60dB		
TOTAL	–85dBm		–89dBm

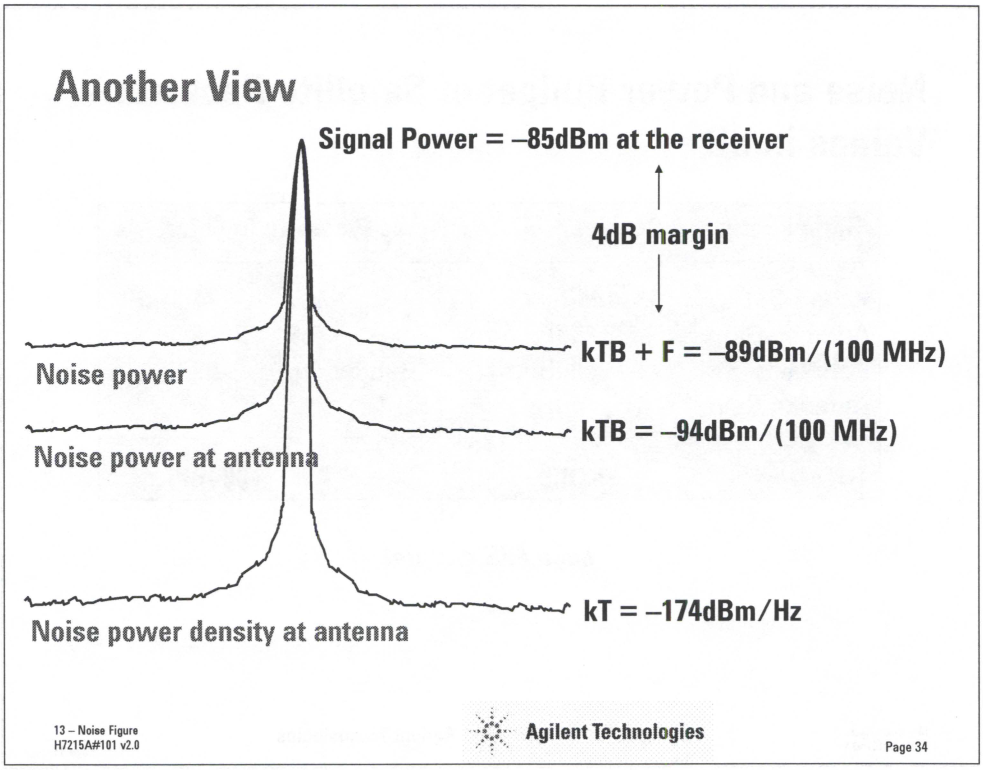
i.e. a 4dB margin!

13 – Noise Figure
H7215A#101 v2.0

Agilent Technologies

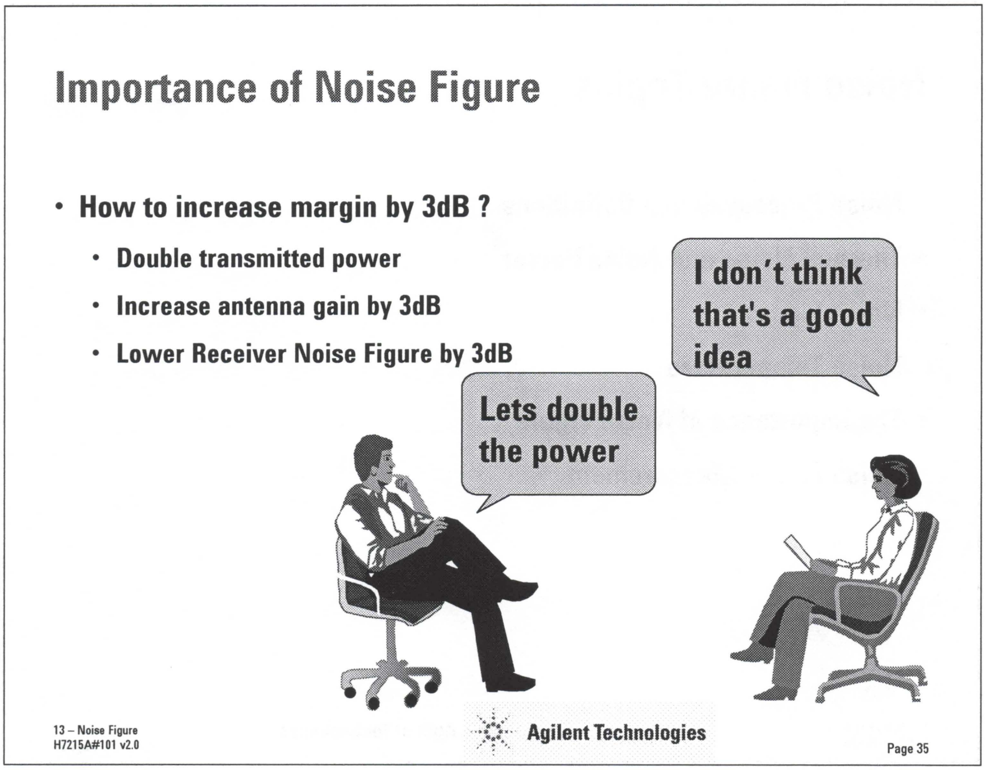
Page 33

This is the power budget for the system, note how only NF and bandwidth are in the noise column. The result is clear; there is only a 4dB instead of the required 6dB S/N ratio (the signal is 37% less than necessary).



One objection that this view of the problem shows is that we have added the noise figure but not the gain of the receiver. The margin calculation does not need this. If the receiver gain was for example 30dB, then 30dB would be added to both the signal and the noise power numbers. The margin would still be 4dB.

If the required margin was indeed 6dB, the system has missed by 2dB. What could one do to improve the situation? How could we increase the margin by 3dB?



It would be very impractical to increase the transmitted power, a new satellite with higher transmitted power is not a practical solution.

In order to increase antenna gain by 3dB the diameter must be increased to double it's area. The satellite antenna can not be changed and improving the ground antenna while possible may not be the cost effective solution.

The best solution, most cost effective, would be to specify a lower noise figure for the receiver.

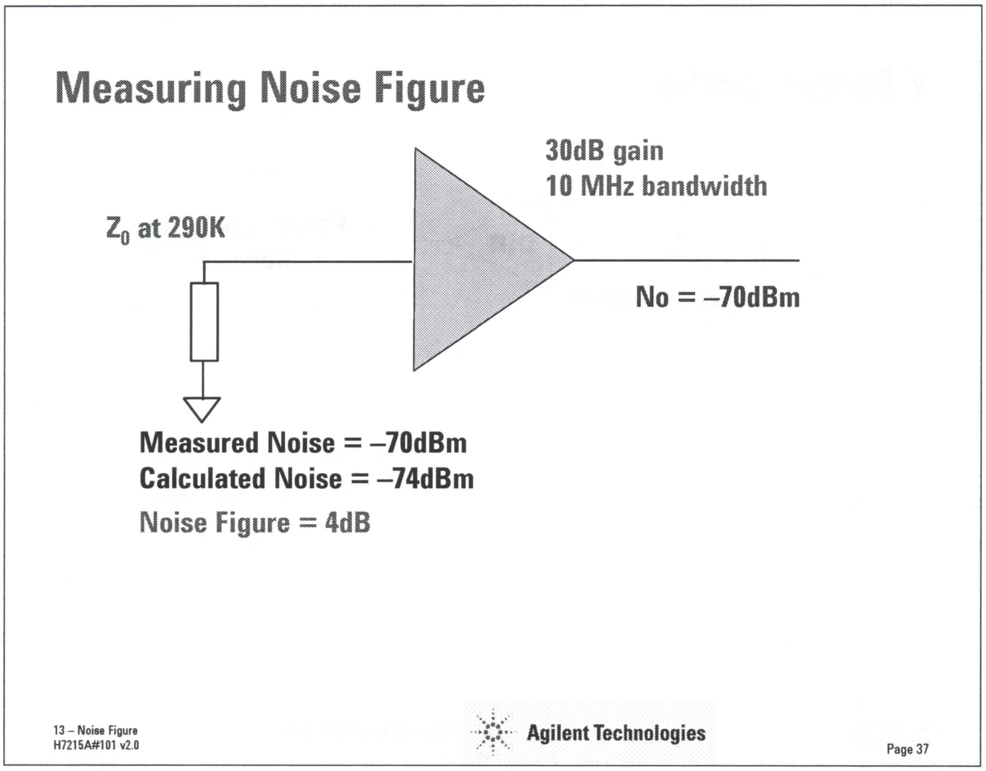
Noise Figure Topics

- Noise Processes and Definitions
- Thermal Noise and Noise Power
- Noise Figure
- Noise Temperature
- The Importance of Noise Figure
- Noise Figure Measurement

13 – Noise Figure
H7215A#101 v2.0

Agilent Technologies

Page 36

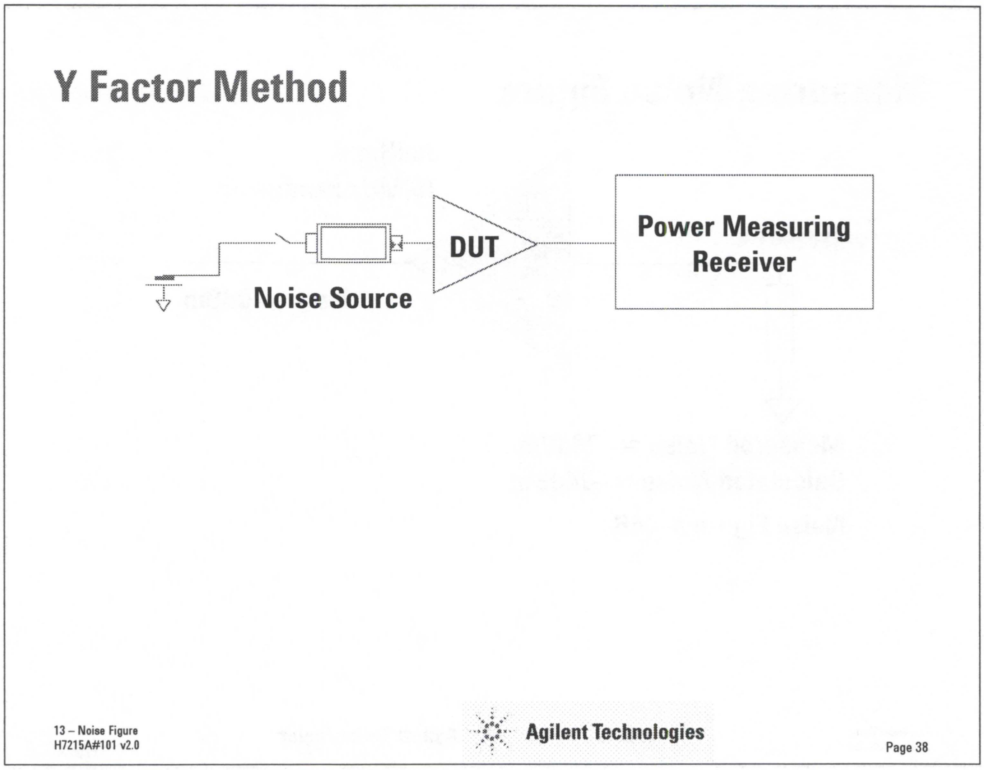


The Simple Way to Measure NF

This is the example that we calculated earlier, it seems simple but in order to measure this we need to:

- Measure the noise bandwidth of the device
- Measure the gain over this BW

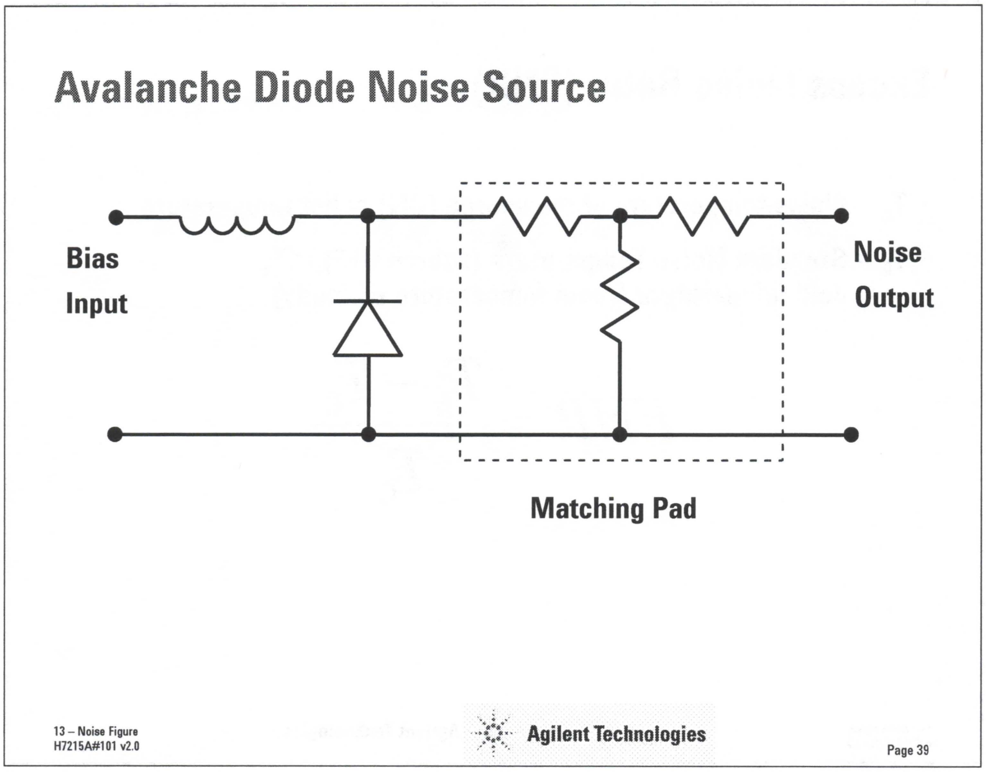
A simple power meter, its sensor being very broadband, would measure the average power over the device BW. In the above example the BW is 10MHz resulting in a measured power of -70dBm, which is at the limit of a diode sensor. A device with lower Noise Figure or Gain or a smaller BW would not be measurable with a diode sensor.



Y Factor Method

The simple method has the problem that available gain and noise bandwidth must be known. The Y factor technique uses a calibrated noise source which is switched on and off, presenting the input of the DUT with two noise power levels. The power measuring device is a selective receiver, that measures the output power levels. The higher level is the equivalent of the noise power from a resistor heated to a temperature significantly above room temperature T0 ; the lower level, a resistor at T0. The Y factor is the linear ratio of these powers. The great advantage of this technique is that prior knowledge of noise BW and gain is not needed. The gain and BW terms cancel, and noise figure is a function of the noise source calibration(ENR, defined next page) and Y.

A variant of this method is used by the Agilent 8970B noise figure meter, by the Agilent 85719A (8590E series based) and by the Agilent 70875 (70000 series based)



Avalanche Diode Noise Source

Rather than heat a resistor, avalanche diodes are used to generate the high-temperature noise.

These noise sources are popular; they are compact , require little power and are stable with time. The the matching pad minimizes the ON to OFF (almost "short" to "open") impedance change presented to the device being tested.

The most common coaxial noise sources such as the Agilent 346B have an Excess Noise Ratio (ENR) of about 15dB. Note that ENR is not an absolute value but rather indicates the ratio of the output noise power when the diode is biased ON to the output noise power when the diode is OFF. The pad in these sources sources may not sufficiently isolate the diode mismatch for some applications, such as transistor characterization. Specially designed sources such as the Agilent 346A with an ENR of about 6dB are used for such measurements.

Excess Noise Ratio (ENR)

- T_h = Noise temperature of the source (ON) = hot temperature
- T_0 = Standard Noise Temperature (source OFF) = T_c = cold temperature (room temperature typically)

$$ENR = \frac{T_h - T_0}{T_0}$$

13 – Noise Figure
H7215A#101 v2.0

Agilent Technologies

Page 40

Excess Noise Ratio

This is ENR expressed in linear terms, this parameter is usually expressed in dB:

$$ENR_{dB} = 10\log\left(\frac{T_h - T_0}{T_0}\right)$$

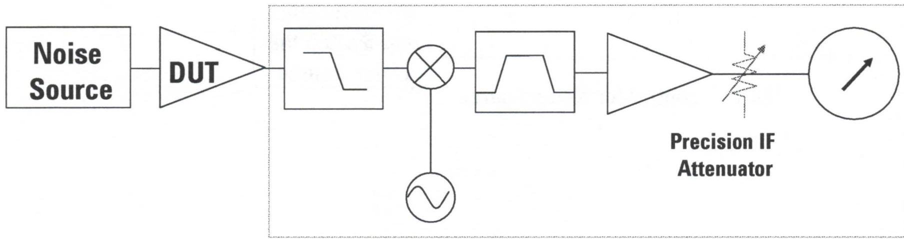
Note: The above definition is used for most practical applications and $T_h = T_a$, the temperature that yields the *available* power spectral density from the source.
The definition used by NIST, however, is when $T_h = T_{ne}$, the temperature that yields power spectral density *delivered to a non-reflective* (and non-emitting) load. The relationship between these temperatures is given by the mismatch loss equation:

$$T_{ne} = T_a(1 - |\Gamma|^2)$$

where Γ is the reflection coefficient of the noise source.

Y Factor Method

- **Where Y = Change in IF attenuator to maintain reading on detector as source switched ON and OFF.**



$$F = 10\log\left[\frac{\left(\frac{T_h}{290}-1\right)-Y\left(\frac{T_c}{290}-1\right)}{Y-1}\right]$$

$$F = ENR - 10\log(Y-1)$$
$$\text{if } T_C = 290K$$

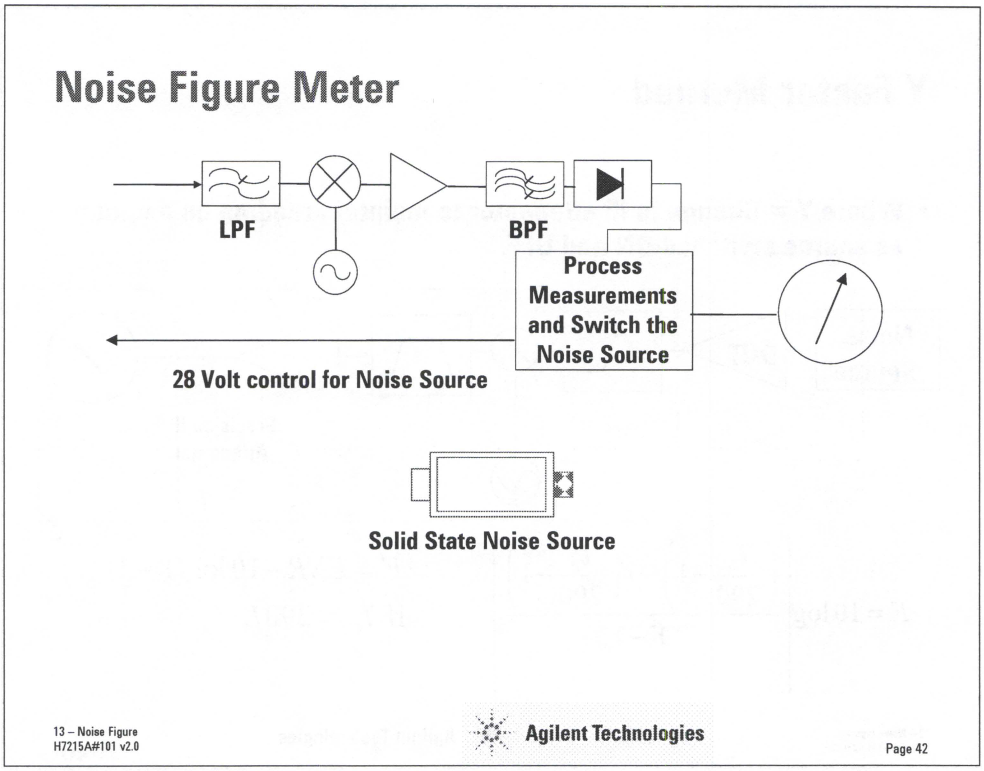
13 – Noise Figure
H7215A#101 v2.0
Agilent Technologies
Page 41

Y Factor Method

For the manual method the operator will adjust the precision attenuator to maintain the same detector reading for the noise source on and off. The change of attenuator setting being the value of Y.

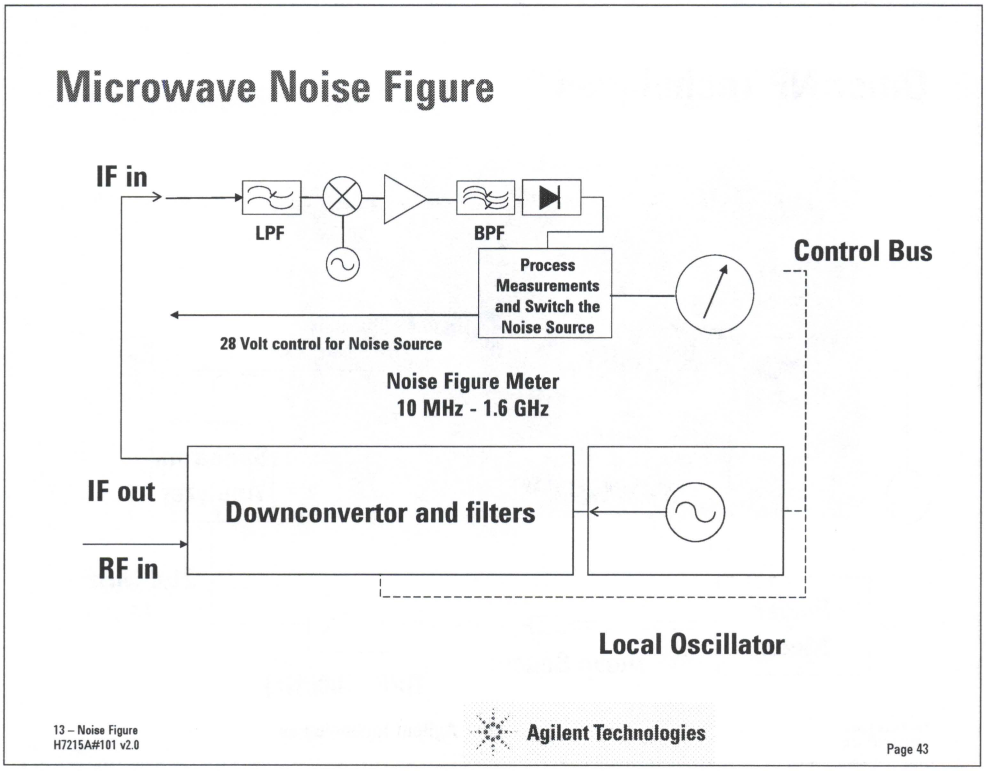
If the change is in dB then the linear equivalent must be calculated for the above expression.

T_h = hot temperature (diode biassed on)
T_c = cold temperature (diode biassed off)



A Noise Figure Meter (e.g. Agilent 8970B)

A noise figure meter implements the Y- factor technique automatically by controlling the noise source and measuring the ON and OFF power levels. It is a multistage receiver.The IF is 4MHz wide. The diagram above shows the major functions as a single stage receiver. Such a system measures Noise Figure and Small Signal Gain from 10 to 1600 MHz. The system will also calibrate itself and correct for second stage effects, the second stage being the noise figure meter itself.

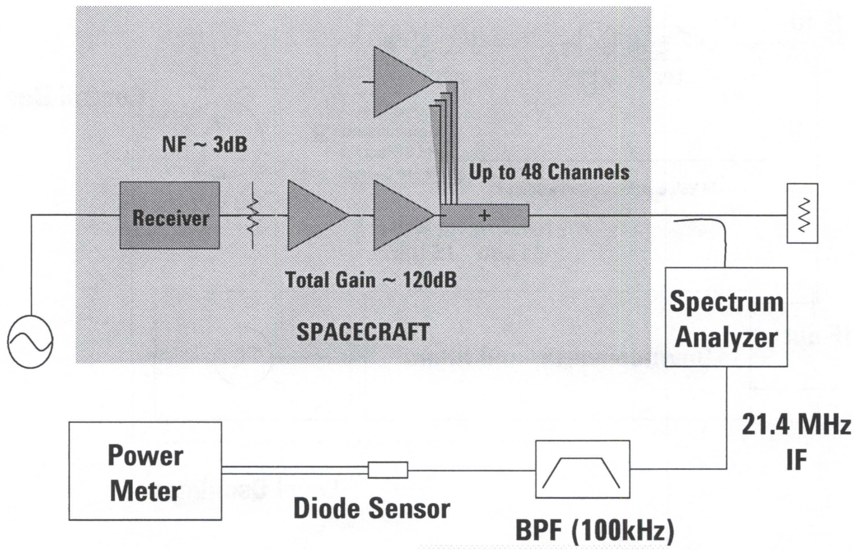


To measure at frequencies higher than the receive range of the noise figure meter, some downconversion is necessary. The diagram above shows such a system. The NFM can issue command codes to the LO and Down Converter in order to make an automatic measurement.

This system can also be used to measure a downconverter (receiver or mixer) by suitable modification of the measurement process.

(See Product Note 8970B/S-2)

Other NF Techniques



13 – Noise Figure
H7215A#101 v2.0

Agilent Technologies

Page 44

A version of an old method is used to test satellites. This method uses a spectrum analyzer and a signal generator to give a calibrated CW signal. A spectrum analyzer always measures signal plus noise, so the correction to S/N must be made.

$$F_{dB} = 10\log\left(\frac{N_a}{kT_0BG}+1\right)$$

From this basic definition a form incorporating S/N may be derived.

$$F_{dB} = 10\log P_{in} - [10\log kT_0B + 10\log(\frac{S}{N})]$$

The measurement procedure uses the spectrum analyzer in zero span to downconvert the carrier + noise to an IF which is measured by a power meter via a BPF. (BW_noise). The carrier frequency is then de-tuned away from the S/A tuned frequency to measure the noise floor N. The S/N is calculated from the S+N and N measurements.

$$\frac{S+N}{N} = \frac{S}{N}+1 \quad \text{when S, and N in linear terms (watts)}$$

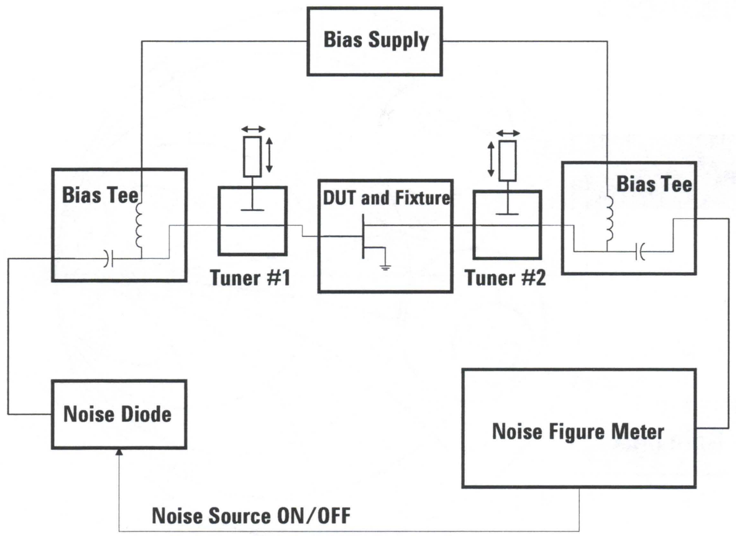
But since S and N are actually measured in dBm's

Then (S-to-N_dB)

$$= 10\log(10^{((S+N)-N)/10} - 1)$$

Now P_{in} is known and the (BW_Noise) may be determined from the BW of the IF BPF's.

Transistor Noise Measurement

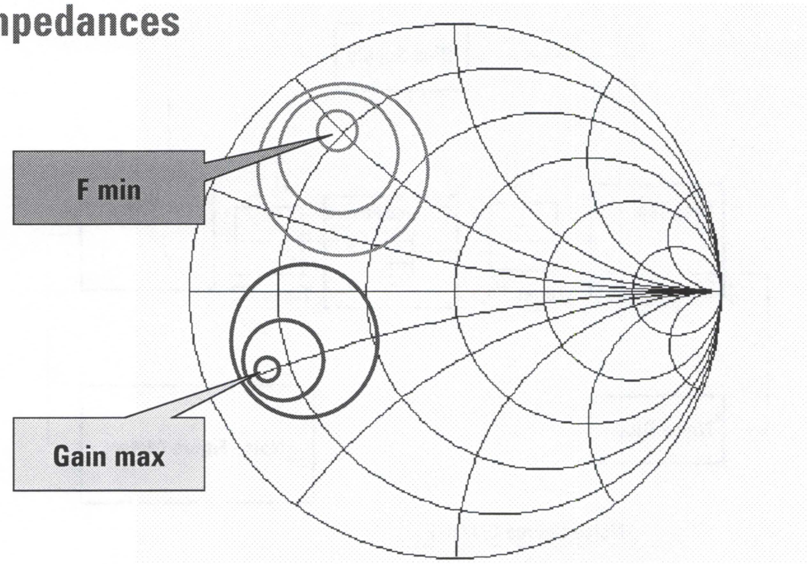


13 – Noise Figure
H7215A#101 v2.0

Agilent Technologies

Page 45

Noise Figure & Gain Circles At Different Source Impedances



Gain Review

- **Insertion Gain**
  - This is what is usually measured Gi = Pd/Pr
  - Pr = power to load when source is connected to load
  - Pd = power to load when device is connected
- **Power Gain**
  - G = Pl/Ps
  - Ps = power to device
  - Pl = power to load from device
- **Available Power Gain**
  - This is what is needed Ga = Pao/Pas
  - Ps = power available from source
  - Pl = power available to load from device

13 – Noise Figure
H7215A#101 v2.0

Agilent Technologies

Page 47

Summary

- Thermal Noise limits communications.
- Noise power = kTB = –174dBm; T=T0, B = 1Hz.
- Noise figure is degradation of S/N.
- Any two port device has a noise figure.
- First gain stage contributes most to noise figure.

13 – Noise Figure
H7215A#101 v2.0

Agilent Technologies

Page 48

Cold Source Method

Most of the methods so far mentioned rely on power measurements which are too slow for some production needs. An alternative technique is one which uses the 8530A microwave receiver, it is called the "cold source" technique because the noise source is used only during calibration. This has been demonstrated to be fast and accurate. The method is applied in some ATE systems and is still evolving.

For more information on this topic please refer to Agilent Technologies Application Note "Fundamentals of RF and Microwave Noise Figure Measurements" AN 57-1.

Phase Noise Measurement Basics
(An Overview)

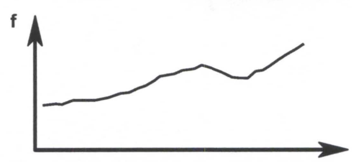
Agilent Technologies

What is Phase Noise?

Signal (frequency) instability:

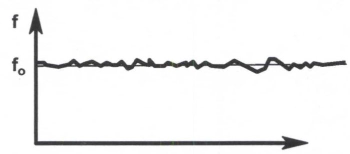
- Large variations are residual FM or drift
- Small variations are phase noise and amplitude noise

Large Frequency Instability



Large variation in frequency
(seconds, minutes, days, months, years)

Small Frequency Instability



Small instantaneous frequency variations
around the nominal frequency

Agilent Technologies

SLIDE 2

What is Phase noise?

There are different ways of characterizing the stability of a signal. This stability can generally be divided into three components:

long-term stability,

short-term stability, and

stability due to environmental shifts.

Long-term frequency stability, or aging, describes the variation in signal frequency that occurs over long periods of time (days, months, or years). Short-term frequency stability, on the other hand, is the variation in signal frequency over time periods of less than a few seconds. The dividing line between short and long-term stability is a function of application. Environmental shifts in the frequency can be due to changes in pressure, temperature, humidity, etc.

For our discussion during this class we will be talking about short term frequency instability. In discussing short term frequency stability, there are two "classes" of frequency variations: non-random (or deterministic) and random. The first, deterministic, are discrete signals which appear as distinct components on our ideal spectrum. These signals, commonly called spurious, can be related to known phenomena in the signal source such as power line frequency, vibration frequencies, or mixer products.

The second type of variation is random in nature, and is commonly

Ideal and Real Signals

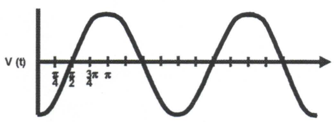
IDEAL SIGNAL

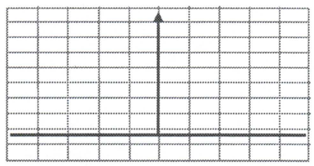
$$V(t) = A_0 \sin(2\pi f_0 t)$$

where

A_0 = Nominal Amplitude

f_0 = Nominal Frequency





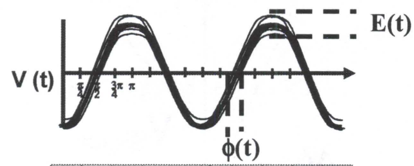
REAL WORLD SIGNAL

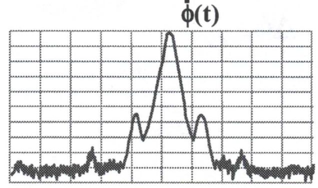
$$V(t) = (A_0 + E(t))\sin(2\pi f_0 t + \phi(t))$$

where

$E(t)$ = Random amplitude flucuations

$\phi(t)$ = Random phase flucuations





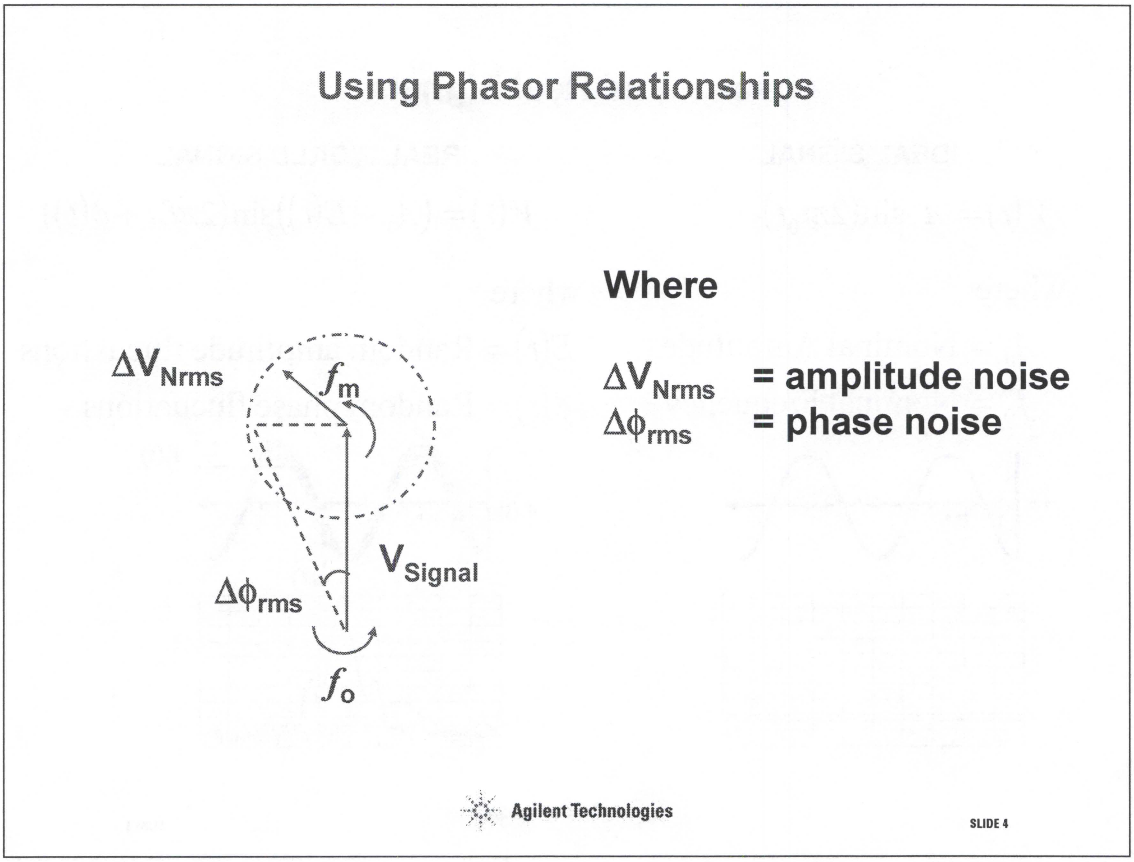
Agilent Technologies

SLIDE 3

Before we get to the formal definitions of phase noise, let's look at a couple of difference signals. The first being an ideal signal (a perfect oscillator). In the frequency domain, this signal is represented by a single spectral line. In the time domain, this signal is a perfect sinusoid waveform.

In the real world however, there is always small, unwanted amplitude and phase fluctuations present on the signal. Notice that frequency fluctuations are actually an added term to the phase angle portion of the time domain equation. Because phase and frequency are related, you can discuss equivalently about unwanted frequency or phase fluctuations.

In the frequency domain, the signal is no longer a discrete spectral line. It is now represented by a spread of spectral lines - both above and below the nominal signal frequency in the form of modulation sidebands due to random amplitude and phase fluctuations.

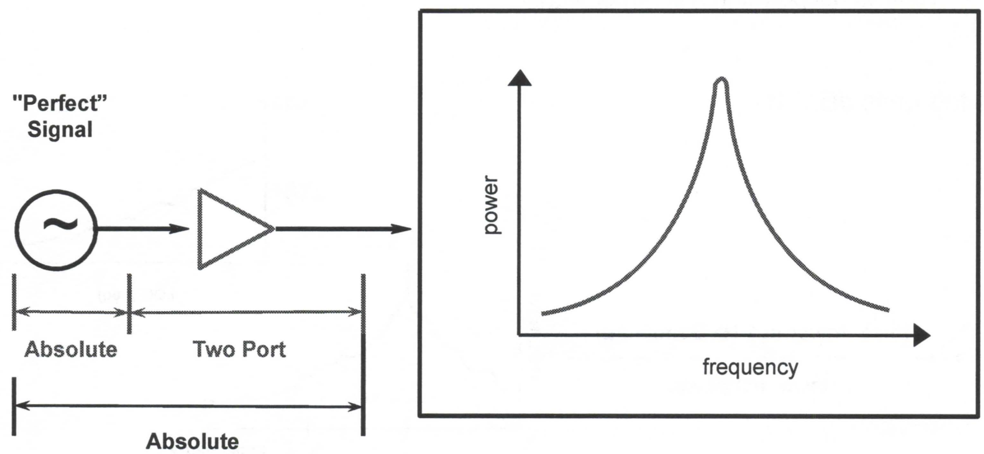


You can also use phasor relationships to describe how amplitude and phase fluctuations affect the nominal signal frequency. Here you have the main signal V_{Signal} with a carrier frequency of f_o. The amplitude and phase noise are a small vector that adds to the main carrier.

Two Port vs. Absolute (Total) Noise

Absolute - Specified on sources or complete system

Two Port - Specified on devices or subsystem



Agilent Technologies

SLIDE 5

Discussions about phase noise can be divided into two topics: the total or absolute noise from an oscillator or system that generates a signal, a device with only one port, and the added or "two-port" noise that is added to a signal as it passes through a device or system.

Absolute noise measurements on the output signal of a system would include the noise that occurs where the signal is generated. This could be a simple oscillator or a complete system (oscillator, amplifiers, filters) that generates a signal.

"Two-port" noise refers to the noise of devices (amplifiers, dividers, mixers, etc). Two-port noise is the noise contributed by a device, regardless of the noise of the reference oscillator used. One way to look at "two-port" noise is how much noise would be added by a device if a perfect signal were input to it.

Example:

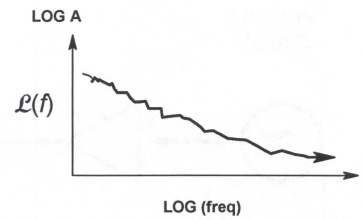
A system, such as a synthesizer, has both two-port and absolute noise associated with it. The reference signal of the synthesizer, comprising an oscillator element, has absolute noise. The synthesizer, comprising of phase-locked loops, multipliers, dividers, etc., have some two-port noise contribution. The integrated system, the synthesizer, also has a value for absolute noise, or all noise present at the output.

Basic Phase Noise Concepts - Unit of Measure

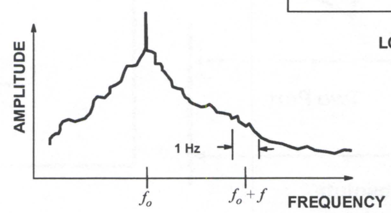
Usually phase noise is quantified as $\mathcal{L}(f)$ and is defined by two measurements.

- Noise power measured in a 1 Hz bandwidth at a frequency *f* from the carrier.
- Total signal power in the carrier signal

Has the units dBc/Hz



$$\mathcal{L}(f) = \frac{\text{Noise Power in 1 Hz Bandwidth}}{\text{Total signal power}}$$



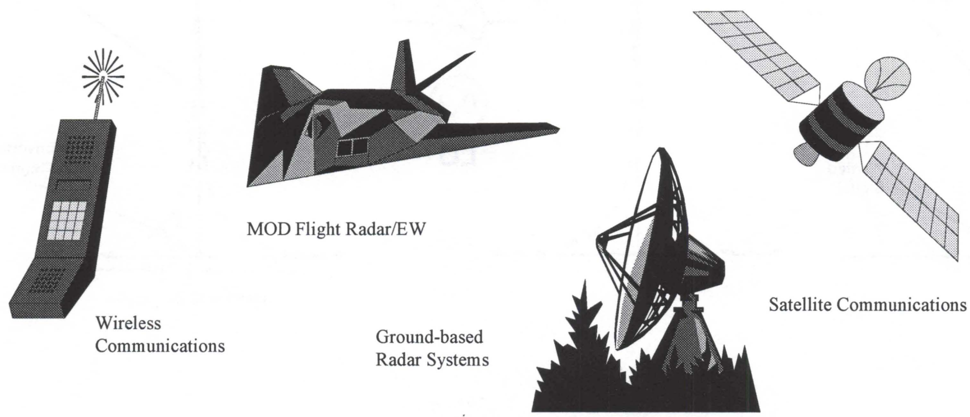
Agilent Technologies

SLIDE 6

Historically, and the typical way of measuring phase noise is as a single sideband phase noise power. This is done by measuring the power in a 1 Hertz bandwidth at a specific offset frequency, f, and dividing it by the total power in the carrier signal. This unit of measure is described as L(f) (pronounced script L(f)) and has the units of dBc/Hz.

Why is Phase Noise Important?

Phase noise is often THE limiting factor in many RF frequency based applications and varies in importance within those applications...



E5500 Phase Noise Measurement
Training Course

Agilent Technologies

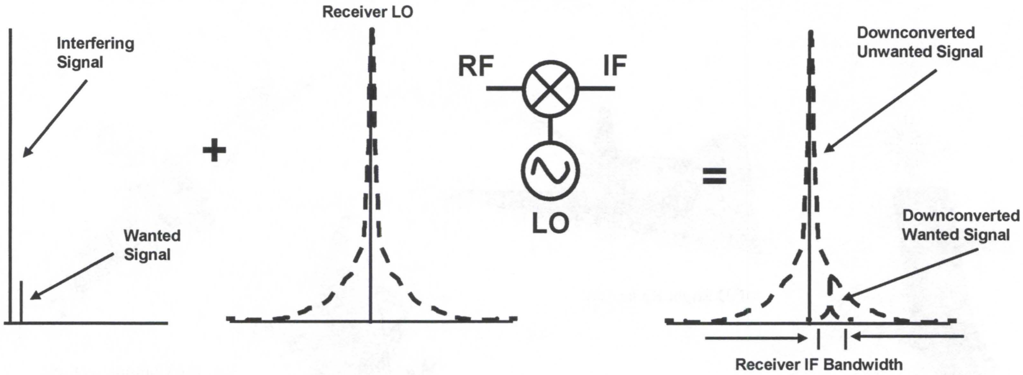
Page 5

The importance of phase noise varies with the application concerned. It is often the limiting factor in many RF based wireless applications and is the limiting factor in many radar systems.

Why is Phase Noise Important?

Adjacent Channel Noise

Local Oscillator phase noise affects Receiver sensitivity:



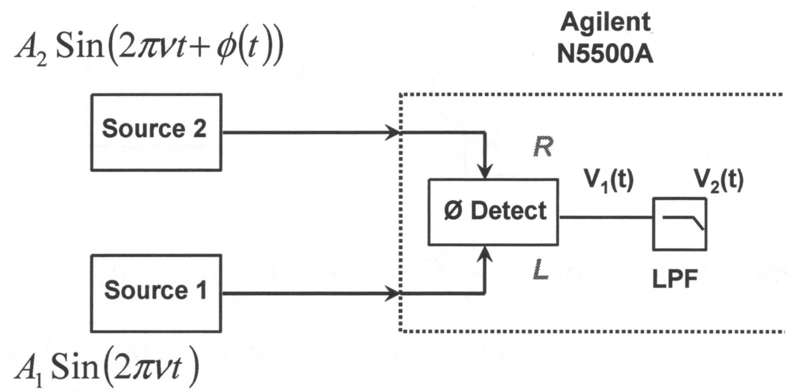
Agilent Technologies

SLIDE 8

In many communication systems, detecting a small signal in the presence of a large one is a major challenge. This challenge is in part related to the phase noise characteristics of the receiver and in particular, the local oscillator within the receiver. This also applies to radar systems as well.

If the two signals were perfect - no noise - then a simple filter could eliminate the unwanted large signal. But in most receivers, the RF signals are mixed with an internal local oscillator to move the signal(s) to a much lower frequency. The LO within the receiver has phase noise side-bands which are transposed to the received signals. Even with channel filters, there is enough noise energy related to the large unwanted signal appearing in the desired signal channel that the desired signal is no longer detectable. Lower phase noise characteristics of the receiver LO would help.

Phase Detector Basics



Agilent Technologies

SLIDE 9

To understand how to measure phase noise, a brief discussion of how the a phase detector works is required.

A double-balanced mixer is used as the phase detector. When two signals at identical frequencies and relative phase difference of 90 degrees are inputted to the phase detector, it becomes a phase demodulator.

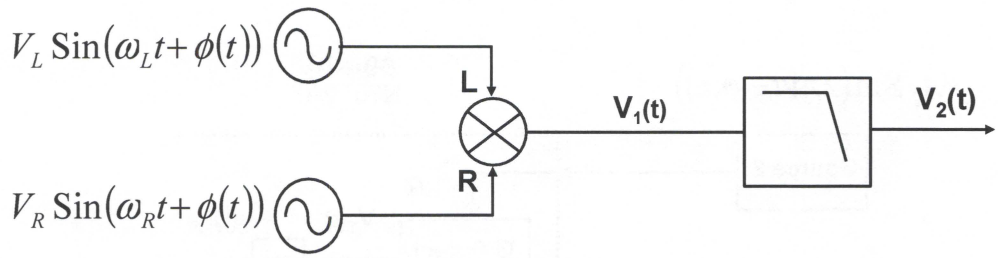
Consider two signals:

1) a perfect sine wave represented by a cosine equation, and

2) a sine wave with phase noise represented by a sine equation (the value of A must be high enough to drive the mixer into its linear operating range).

These two signals are input into the reference and test ports of the phase detector. The phase detector is within the Agilent N5500A phase noise set.

Double Balanced Mixer Used As A Phase Detector



$$\text{Using} \quad \operatorname{Cos} x \cdot \operatorname{Cos} y = \frac{1}{2}\{\operatorname{Cos}(x+y)+\operatorname{Cos}(x-y)\}$$

where K_L = mixer efficiency

$$V_1(t) = K_L V_R \operatorname{Cos}((\omega_R - \omega_L)t + \phi(t)) + K_L V_R \operatorname{Cos}((\omega_R + \omega_L)t + \phi(t)) + \quad \ldots$$

Odd harmonics of ω_L are present due to the square wave switching function

$$V_2(t) = K_L V_R \operatorname{Cos}((\omega_R - \omega_L)t + \phi(t))$$

Agilent Technologies

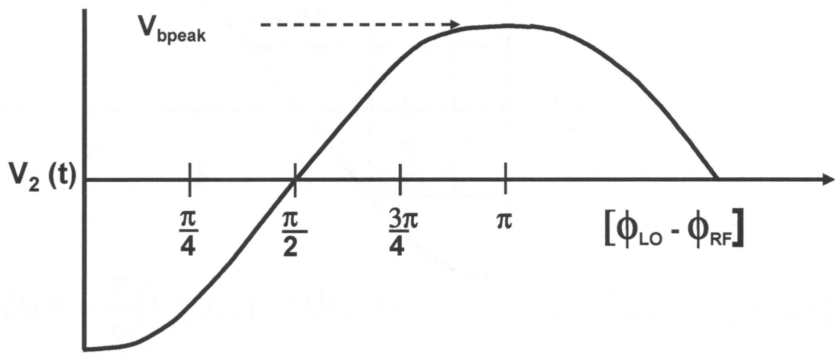
SLIDE 10

Let us assume for the moment that these two signals are at different frequencies and that one of them is driving the mixer with enough power to ensure that the output IF power is independent of the LO port power.

The output IF ($V_1(t)$) power is scaled by the mixer conversion loss and comprises sum and difference signals of the two input frequencies.

The sum signals are then filtered out by the low pass filter and the difference signal $V_2(t)$ is fed to the LNA and baseband processing.

Double Balanced Mixer Used As A Phase Detector



$$V_2(t) = K_L V_R \operatorname{Cos}((\omega_R - \omega_L)t + \phi(t))$$

$$\therefore \omega_R = \omega_L$$

$$V_2(t) = K_L V_R \operatorname{Cos}(\phi(t))$$

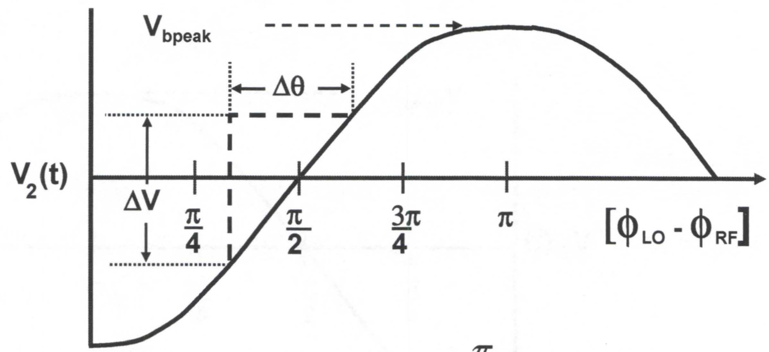
Agilent Technologies

SLIDE 11

Consider the situation that will occur if the two signals are at the same frequency.

V(t) is now time invariant and will have a voltage that is a cosine function whose peak value is the peak value of the beatnote that would occur if the frequencies were NOT at the same frequency. The value of V(t) is zero when the two signals (at the same frequency) are 90 degrees out of phase with one another (i.e. in phase quadrature). This is the point on the phase detector's characteristic that we wish to operate. What is required is the slope of this curve at this point as that will yield the Phase Detector Constant, K_L.

Double Balanced Mixer Used As A Phase Detector



$$V_2(t) = V_{bpeak}\,\mathrm{Cos}(\theta(t))$$

$$\Delta V_2(t) = \pm V_{bpeak}\,\mathrm{Sin}\,\Delta\theta(t)$$

$$\Delta V_2(t) = \pm V_{bpeak}\Delta\theta(t)$$

$$\Delta V_2(t) = K_{\phi peak}\Delta\theta(t)$$

Let $\theta(t) = (2N+1)\dfrac{\pi}{2} + \Delta\theta(t)$

For $\Delta\theta_{peak}(t) \le 0.2$ radians

$$\mathrm{Sin}\,\Delta\theta(t) \approx \Delta\theta(t)$$

K_ϕ = phase detector constant (volts/radian)
= V_{bpeak}

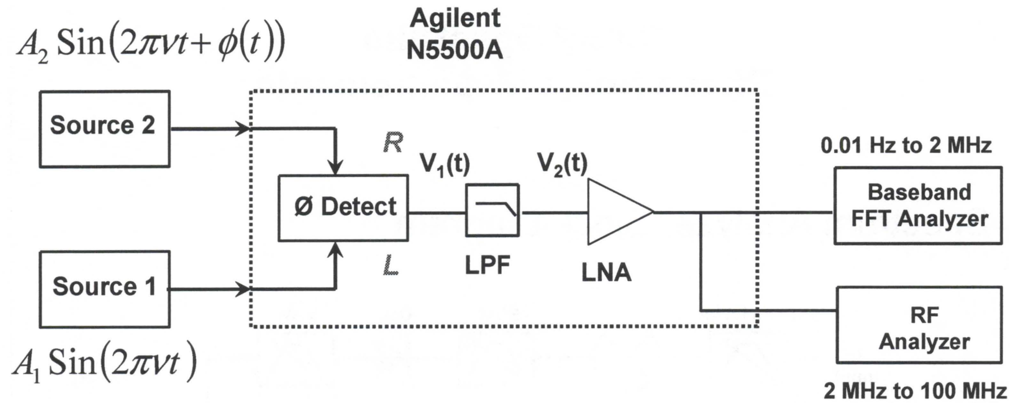
Agilent Technologies

SLIDE 12

Assuming that the phase deviations are small on either side of the quadrature point. It is simple to show that the phase detector constant K_ϕ is numerically equal to the peak value of the beat note.

The E5500 system uses this relationship as part of the calibration techniques by forcing the two signals apart in frequency, measuring the value of this beatnote and extracting the value of K_ϕ.

Measuring the Baseband Signal



Low Noise Amplifier (LNA) overcomes sensitivity of spectrum analyzer

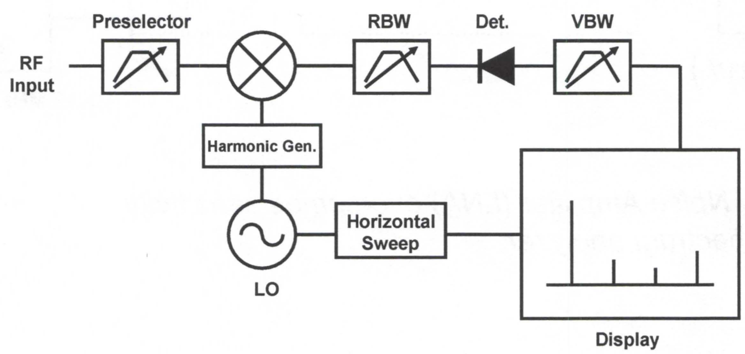
Agilent Technologies

SLIDE 13

After the filter, the signal is amplified to overcome the noise floor of the FFT analyzer and RF spectrum analyzer.

Direct Spectrum
Phase Noise Measurements

Spectrum Analyzer Block Diagram



Agilent Technologies

SLIDE 14

At this point, let's go back to our spectrum analyzer (direct spectrum measurement) and understand exactly why it doesn't have the performance we need for phase noise measurements.

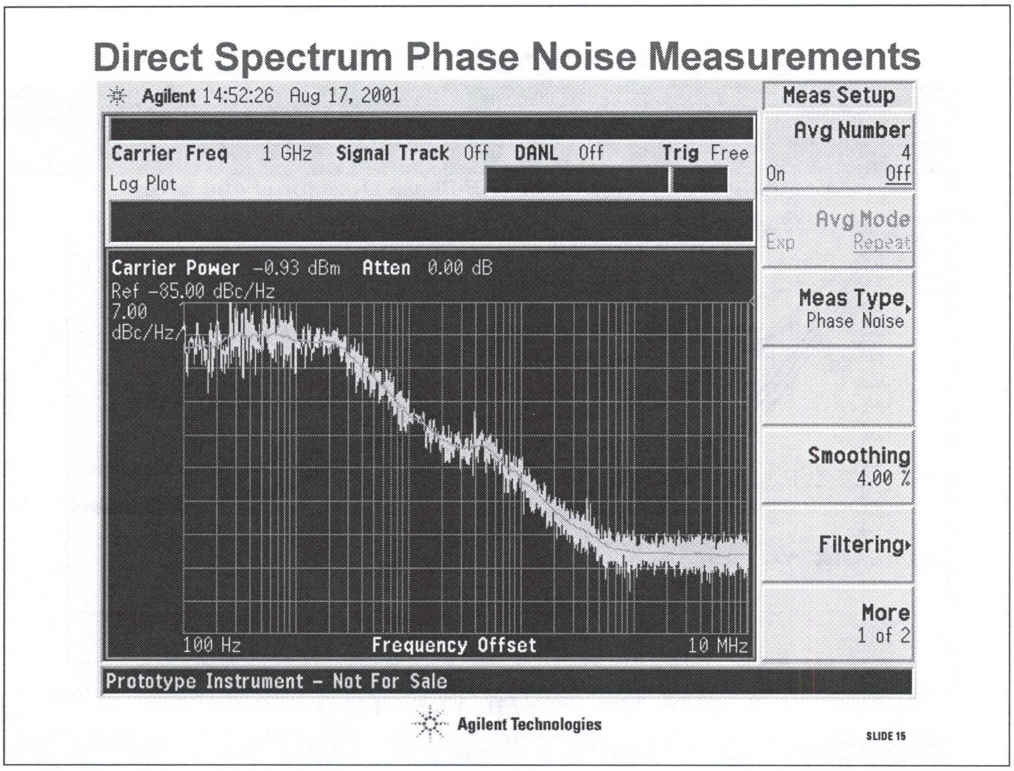
The front end of the SA is basically made up of a mixer and an LO. We take in the signal and downconvert it once or twice to an IF we can more easily handle. For RF spectrum analyzers, the LO tunes over the frequency range we want to measure. For microwave spectrum analyzers, the LO may drive directly into some kind of harmonic generator and just tune over a limited frequency. We mix down the signal over a frequency, say 2-4 GHz, then relock to a harmonic and mix down over another range, say 4-8 GHz.

As the LO is being driven by the horizontal sweep generator, the corresponding amplitude being seen by the diode detector is displayed on the screen.

The preselector is a tuning filter that helps to keep the mixer from being a broadband front end for the spectrum analyzer. This helps to keep images from making their way down the detection path. The signal is then filtered, detected and displayed.

Although most of our spectrum analyzers have a synthesized LO, their performance is somewhat lacking.

For phase noise measurements, a perfect spectrum analyzer might have an 8663 for an LO.



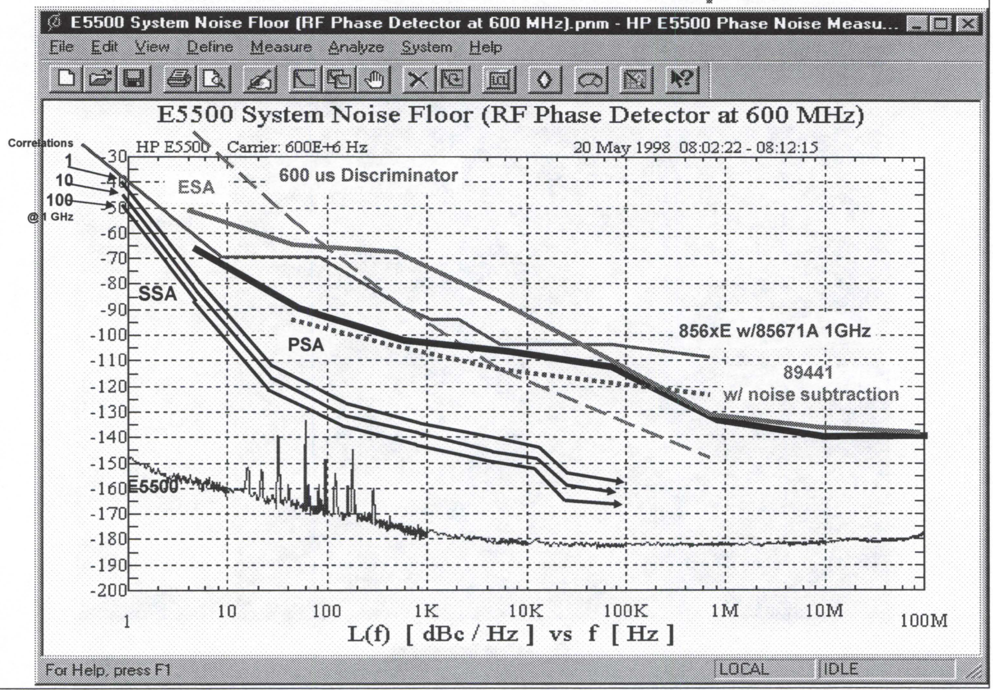
Agilent Technologies sells Phase Noise measurement capabilities with our spectrum analyzers. So we must be able to get a good measurement, right?

So, it comes down to answering the question, "how good is the LO of a Spectrum Analyzer?" The above plot shows the performance of the Phase Noise Measurement Personality, 85671A, that can be installed on the 856xE series Spectrum Analyzers. Take a look at this plot and compare it to two slides ago for the 8662/3. How do they compare?

To help keep the cost of the SA down, the LO is typically not the high performance source we think it is or should be.

The analyzer's dynamic range (i.e. simultaneous max to minimum signal range) is insufficient for noise >100 dB less than the carrier. Also, the analyzer's detector circuitry responds to both the phase and amplitude modulation of a signal.

Phase Noise Performance Comparison



This graph compares the performance of the current spectrum analyzers performance and should give you a good idea of what you're phase noise measurement techniques might be today using them. The E5500 system plot is provided to show the best possible performance in phase noise measurement.

Choosing the Correct Reference Source

- Since the phase detector maintains the phase noise through the downconversion process, you would like to know the performance of the sources, and you want that source to have better performance than the source under test.

- The performance is driven by price

- Reference source does not have to be Agilent Technologies

- Reference source does hot have to be tunable if the UUT is tunable

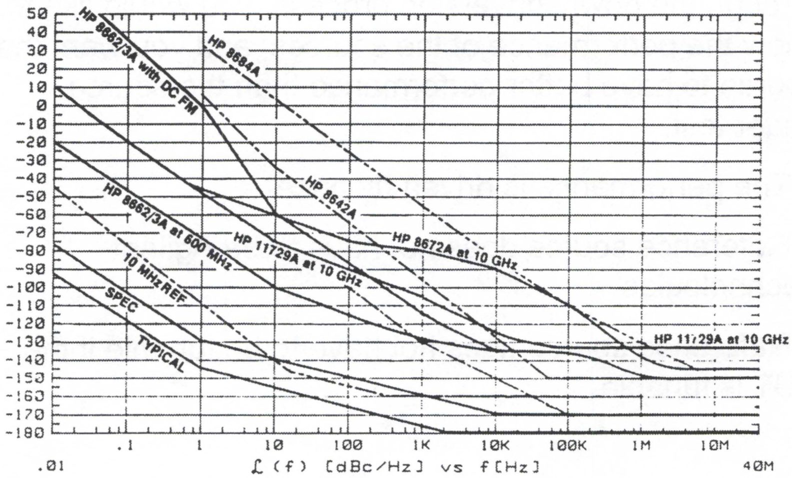
Agilent Technologies

SLIDE 17

One of the most important pieces (and possibly the most expensive) of the phase noise measurement is the reference source. As can be expected, the performance is driven by price. The only requirement of the reference source is that it is well known for phase noise. It does not necessarily have to be an Agilent Technologies source, although we probably know our sources best.

Either the reference source, or the UUT can be tuned to maintain the PLL. Hence the reference source does not necessarily need to be tuned. An example of this is where you have two sources under test with approximately the same performance, and either one can be tuned.

Choosing the Correct Reference Source



Agilent Technologies

SLIDE 18

The above plot shows some references with respect to the noise floor of the 3048A.

Recommended Reference Sources

Source	Tuning Methods			Comments
	DCFM	EFC	External Timebase	
8642A/B	Yes	No	No	Excellent Spectral Purity > 10 kHz
8643A	Yes	No	No	Excellent Spectral Purity > 10 kHz, Wideband Tuning
8644B	Yes	No	No	Excellent Spectral Purity > 10 kHz, Wideband Tuning
8657A	Yes	No	No	Excellent Spectral Purity > 10 kHz, Wideband Tuning
8662A	Yes	Yes	Yes	Excellent Spectral Purity < 5 kHz, Narrow Tuning Bandwidth
8663A	Yes	Yes	Yes	Excellent Spectral Purity < 5 kHz, Narrow Tuning Bandwidth
8664A	Yes	No	No	Excellent Spectral Purity > 10 kHz, Wideband Tuning
8665A	Yes	No	No	Excellent Spectral Purity > 10 kHz, Wideband Tuning
8665B	Yes	No	No	Excellent Spectral Purity > 10 kHz, Wideband Tuning
E44xx	Yes	No	No	Excellent Spectral Purity > 10 kHz, Wideband Tuning, 16 models
203x	Yes	No	No	Marconi Source Support
204x	Yes	No	No	Marconi Source Support

Agilent Technologies

SLIDE 19

There are several tunable, RF reference sources available for automatic use with the Agilent E5500 series. Included are the Agilent 8662A, Agilent 8663A, Agilent 8642A/B, Agilent 8643A, Agilent 8644B, Agilent 8664A, Agilent 8665A/B, Agilent ESG series, and Agilent 8657A/B. Though there are other signal generators that offer tunability, only these generators will be discussed since they offer the best noise performance and/or DCFM range and they can be controlled by the system software through their respective software ACM (asset control module).

This above list are sources that can be controlled by the E5500 software. If you wish to use another source, then you must use the E5500 software with a manual reference source. The E5500 software will prompt you for the required actions.

All the source use frequency modulation (DCFM) to lock the reference source to the source under test. Only two sources, the 8662A and 8663A, offer other techniques for for changing the frequency of the reference source.

Summary

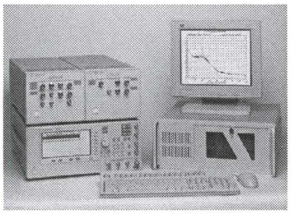
• Phase Noise is a signal's short-term stability caused by non-random and random frequency variations.

• Phase Noise systems measure a single sideband of a signal in a 1Hz bandwidth at a specific offset frequency, and divides it by the total power in the carrier.

• It is often the limiting factor in many RF based wireless applications and radar systems.

• The accuracy of a phase noise measurement with a spectrum analyzer is determined by analyzer's L.O. performance.

• When choosing a reference source, you want that source to have better phase noise performance than the source under test.



Agilent Technologies

SLIDE 20

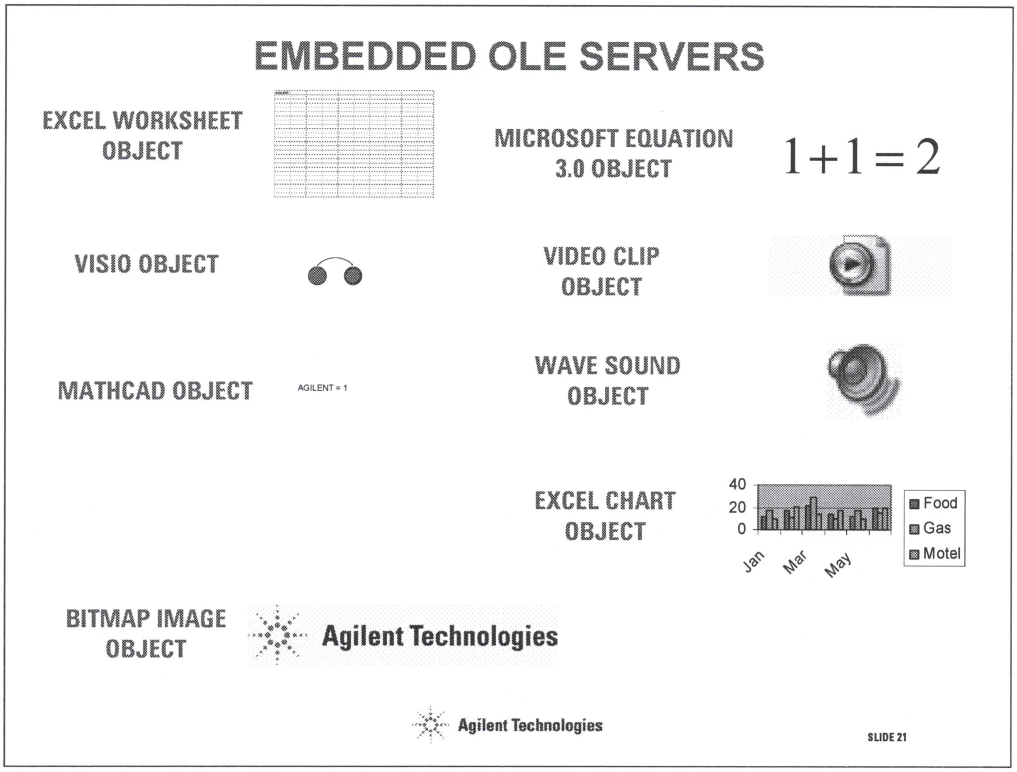
Some key points to remember about phase noise are:

What is phase noise?

Why do we care about it?

How to we measure it?

What equipment is best for your application?



THIS SLIDE CONTAINS OLE OBJECTS AND CAUSES EMBEDDING OF THE RESPECTIVE OLE SERVERS IN THE TEMPLATE.

LEAVE THIS SLIDE AT THE VERY END OF THE PRESENTATION WHICH SHOULD FORCE THE OLE SERVERS TO REMAIN EMBEDDED IN YOUR PRESENTATION.

PLEASE EMBED ANY ADDITIONAL OLE SERVERS YOUR PRESENTATION MAY REQUIRE.

www.agilent.com
800 593 6632

Education and Training

Whether you need to learn fundamental principles, advanced skills and techniques, or new technologies, expand your measurement expertise with our industry and technology specific training.

Spark your educational insight @

www.agilent.com/find/tmeducation

Bill Hewlett and Dave Packard founders of the Hewlett-Packard company and heritage of Agilent Technologies.

Printed in USA November 1, 2001

Agilent Technologies